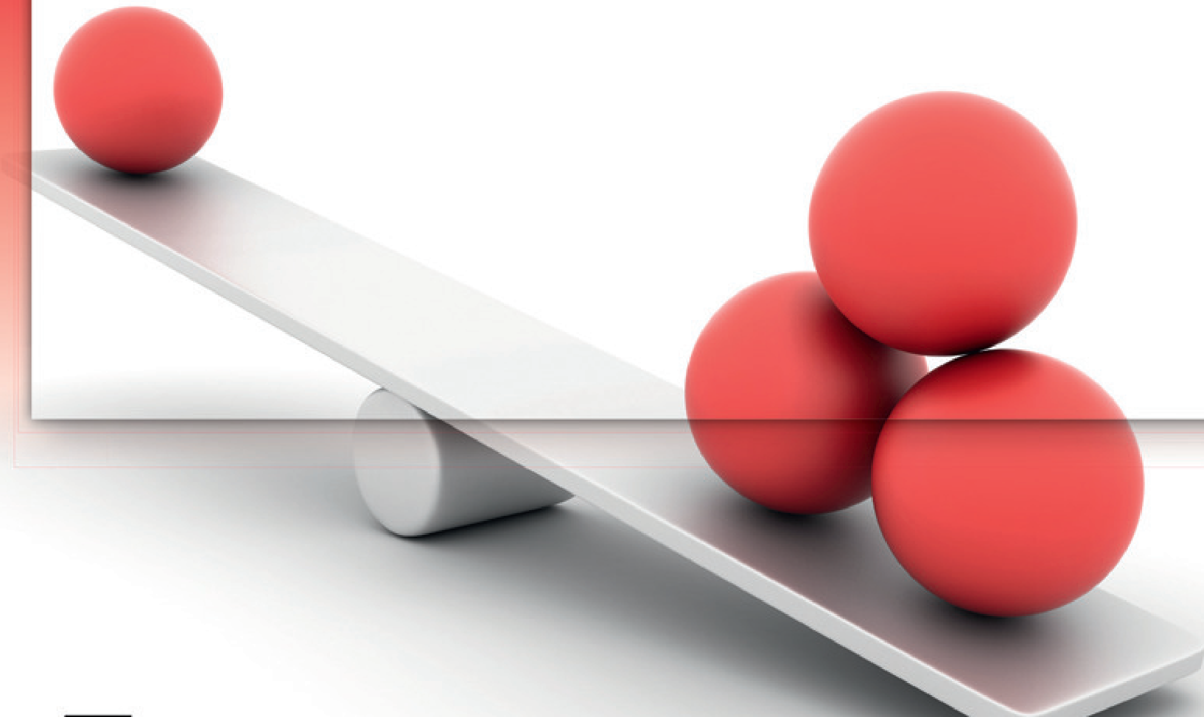
DE GRUYTER

Shigeru Furuichi, Hamid Reza Moradi

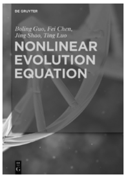
ADVANCES IN MATHEMATICAL INEQUALITIES



Shigeru Furuichi, Hamid Reza Moradi

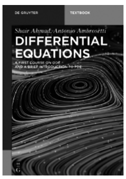
Advances in Mathematical Inequalities

Also of Interest

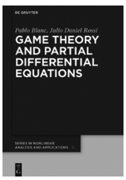


Nonlinear Evolution Equation
Guo, Boling / Chen, Fei / Shao, Jing / Luo, Ting, *to be published 2020*
ISBN 978-3-11-062467-0, e-ISBN (PDF) 978-3-11-061478-7,
e-ISBN (EPUB) 978-3-11-061547-0

Space-Time Methods
Applications to Partial Differential Equations
Ed. by Langer, Ulrich / Steinbach, Olaf, *2019*
ISBN 978-3-11-054787-0, e-ISBN (PDF) 978-3-11-054848-8,
e-ISBN (EPUB) 978-3-11-054799-3



Differential Equations
A first course on ODE and a brief introduction to PDE
Ahmad, Shair / Ambrosetti, Antonio, *2019*
ISBN 978-3-11-065003-7, e-ISBN (PDF) 978-3-11-065286-4,
e-ISBN (EPUB) 978-3-11-065008-2



Game Theory and Partial Differential Equations
Blanc, Pablo / Rossi, Julio Daniel, *2019*
ISBN 978-3-11-061925-6, e-ISBN (PDF) 978-3-11-062179-2,
e-ISBN (EPUB) 978-3-11-061932-4

Shigeru Furuichi, Hamid Reza Moradi

Advances in Mathematical Inequalities

DE GRUYTER

Mathematics Subject Classification 2010
Primary: 47A63, 47A60, 46L05, 47A30, 47A12; Secondary: 26D15, 15A60, 94A17, 15A39, 15A45, 26D07, 47B15, 47A64

Authors
Prof. Shigeru Furuichi
3-25-40 Sakurajousui
156-8550 Tokyo
Japan
furuichi@chs.nihon-u.ac.jp

Dr. Hamid Reza Moradi
Department of Mathematics
Payame Noor University
19395-4697 P. O. Box
Tehran
Iran
hrmoradi@mshdiau.ac.ir

ISBN 978-3-11-064343-5
e-ISBN (PDF) 978-3-11-064347-3
e-ISBN (EPUB) 978-3-11-064364-0

Library of Congress Control Number: 2019955079

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Cover image: porcorex / E+ / gettyimage.de
Typesetting: VTeX UAB, Lithuania
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Preface

In mathematics, the word *inequality* means a disparity between two quantities, which is used to reflect the correlation between two objects. An *inequality* implies that two amounts are not equal. In the 19th century and with the emergence of calculus, the touch of the inequalities and its role increasingly became essential.

In modern mathematics, inequalities play significant roles in almost all fields of mathematics. Several applications of inequalities are found in linear operators, partial differential equations, nonlinear analysis, approximation theory, optimization theory, numerical analysis, probability theory, statistics, and other fields.

The book may be used by researchers in different branches of mathematical and functional analysis, where the theory of Hilbert spaces is of relevance. Since it is self-contained and all the results are entirely proved, the work may also be used by graduate students interested in theory of inequalities and its applications.

For the sake of completeness, all the results presented are wholly proved, and the original references where they have been first obtained are mentioned.

Finally, we would like to thank to Dr. Y. Tsurumi, and Dr. F.-C. Mitroi-Symeonidis for careful checking to improve our manuscript. In addition, this book will not be published without cooperation of all coauthors in our papers. So we would like to thank to all previous coauthors. We also would like to thank to our family since they gave us time to write this book with a smile.

https://doi.org/10.1515/9783110643473-201

Contents

1 Introduction and preliminaries

1.1 Introduction

Mathematical inequalities are applied to many fields on natural science and engineering technology. Some researchers in such fields may find the inequalities from books or papers to solve their problems, while others may discover new inequalities if they do not find them in the literature. For almost all researchers, mathematical inequalities are just tools to solve their problems or improve their results. Such motivations are quite natural to develop their studies, just because of the beauty of their form. On the other hand, mathematicians like to discover new inequalities. They often do not think of its possible applications. A few years ago, I was asked for a proof of the following beautiful inequality by my acquaintance from Japan, when I was in Budapest for a conference. He asked me whether the following inequality is already known. He needed its proof to establish the nonnegativity of thermodynamic entropy for total systems.

$$x > y > z > 0 \quad \Rightarrow \quad x^z y^x z^y > x^y y^z z^x.$$

I gave a proof for his question as follow. Since the representing function $(t-1)/\log t$ is monotone increasing for $t > 0$, we have $(u-1)/\log u > (t-1)/\log t \Leftrightarrow t^{u-1} > u^{t-1}$ for $u > t > 1$. Putting $st = u$, we have $s^{\frac{1}{t}} t^s > st$ for $s, t > 1$. Putting again $t = y/z > 1$, $s = x/y > 1$, we have $(y/z)^{x-y} > (x/y)^{y-z} \Leftrightarrow y^{x-y+y-z} > x^{y-z} z^{x-y} \Leftrightarrow x^z y^x z^y > x^y y^z z^x$.

He also asked this question to several Japanese mathematicians and he reported several proofs during my stay in Budapest. Every proof was different. I enjoyed his report and guessed that so did the other colleagues. A few mathematicians found and proved its extensions to n-variables. After my return, I checked my books on mathematical inequalities, but I could not find the literature. Several mathematicians also could not find the original source of this inequality. Nevertheless, it is natural to consider that this inequality is already known since it is very simple and beautiful in its form. Half a year later, one of ours (H. Ohtsuka) found that this inequality was posed by M. S. Klamkin and L. A. Shepp in [132] in a more general case (n-variables):

$$x_n \geq x_{n-1} \geq \cdots x_2 \geq x_1 \geq 0 \Rightarrow x_1^{x_2} x_2^{x_3} \cdots x_n^{x_1} \geq x_2^{x_1} x_3^{x_2} \cdots x_1^{x_n}, \quad (n \geq 3)$$
$$\text{with equality holding only if } n-1 \text{ of numbers are equal.}$$

We leave the solution of this question to the readers. Mathematicians studying inequalities are often obsessed with a beautiful inequality. Within just one week after the question, several mathematicians gave different proofs of it to each other.

We think the advances of mathematical inequalities may depend on finding applications to other fields and/or interests in inequalities themselves. Here, we present our recent advances on mathematical inequalities. Almost all results were established within the last few years. However, we have added less recent results since they may

https://doi.org/10.1515/9783110643473-001

be interesting for the readers. Some results are obtained to develop the study on entropy theory as applications. There are many good books on mathematical inequalities [30, 31, 35, 37, 98, 106, 241]. Our book cannot cover all areas. We focus on our recent results. In Chapter 2, refinements and reverses of Young-type inequality are given. Other inequalities related to means are given in Chapter 3. The results in Chapter 4 are not recent results, but the unitarily invariant norm inequality and trace inequality are important and interesting for the study of mathematical inequalities. In Chapter 5, we give the recent results on the developments of Jensen-type inequalities with convex theory. In Chapter 6, we summarize recent interesting results on reverses of famous known inequalities. Chapter 7 is also one of our main topics and is devoted to results on estimations of bounds for the relative operator entropy. In Chapter 8, we treat new results on the Kantorovich inequality and uncertainty relation in quantum mechanics as applications of trace inequalities. The results in Section 8.2 are not really new but we believe that the mathematical inequalities are important to develop such a subject. We have tried to organize each chapter by independent contents. However, it is inevitable that some results are connected to others in this wide field. So, each chapter is not always independent.

Almost all of our results are given by operator/matrix inequalities but some of them are given by scalar/numerical inequalities. In addition, we think that the scalar/numerical inequalities are fundamental in almost every case of this book.

1.2 Preliminaries

The purpose of the first chapter is to give basic definitions, some background information and motivations for the problems treated in the later chapters.

Definition 1.2.1. Given a **vector space** $\mathcal{A}$ over the set $\mathbb{C}$ of complex numbers, a norm on $\mathcal{A}$ is a nonnegative-valued scalar function $\|\cdot\| : \mathcal{A} \to [0,\infty)$ with the following properties for all $\alpha \in \mathbb{C}$ and all $a, b \in \mathcal{A}$:
(i) $\|a+b\| \le \|a\| + \|b\|$ (triangle inequality);
(ii) $\|\alpha a\| = |\alpha|\|a\|$;
(iii) $\|a\| = 0$ if and only if $a = 0$.

Definition 1.2.2. A **Banach space** is a vector space $\mathcal{A}$ over the field $\mathbb{C}$ of complex numbers, which is equipped with a norm and which is complete with respect to that norm, that is to say, for every Cauchy sequence $\{a_n\}$ in $\mathcal{A}$, there exists an element a in $\mathcal{A}$ such that $\lim\limits_{n\to\infty} a_n = a$.

Definition 1.2.3. Let $\mathcal{A}$ be a vector space over the field $\mathbb{C}$ of complex numbers equipped with an additional binary operation from $\mathcal{A} \times \mathcal{A}$ to $\mathcal{A}$, denoted here by "$\cdot$" (i. e., if a and b are any two elements of $\mathcal{A}$, $a \cdot b$ is the product of a and b). Then $\mathcal{A}$ is an algebra over $\mathbb{C}$ if the following identities hold for all $a, b, c \in \mathcal{A}$ and $\alpha, \beta \in \mathbb{C}$:

(i) $(a + b) \cdot c = a \cdot c + b \cdot c$;
(ii) $a \cdot (b + c) = a \cdot b + a \cdot c$;
(iii) $(\alpha a) \cdot (\beta b) = (\alpha\beta)(a \cdot b)$;
(iv) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

If $\mathcal{A}$ admits a unit 1 ($a \cdot 1 = 1 \cdot a = a$ for all $a \in \mathcal{A}$), we say that $\mathcal{A}$ is a unital algebra. (In the sequel, we often omit the symbol "$\cdot$".)

Definition 1.2.4. A **normed algebra** $\mathcal{A}$ is an algebra over $\mathbb{C}$ or $\mathbb{R}$ which has a submultiplicative norm:

$$\|a \cdot b\| \le \|a\|\|b\| \quad \text{for all } a, b \in \mathcal{A}.$$

A complete normed algebra is called a **Banach algebra**.

Definition 1.2.5. An **involution** on an algebra $\mathcal{A}$ is a map $a \mapsto a^*$ on $\mathcal{A}$, such that for all $a, b \in \mathcal{A}$ and $\alpha \in \mathbb{C}$:
(i) $a^{**} = (a^*)^* = a$;
(ii) $(a + b)^* = a^* + b^*$;
(iii) $(a \cdot b)^* = b^* \cdot a^*$;
(iv) $(\alpha a)^* = \overline{\alpha} a^*$.

The pair $(\mathcal{A}, *)$ is called a $*$**-algebra**.

Definition 1.2.6. A **Banach** $*$**-algebra** $\mathcal{A}$ is an $*$-algebra together with a complete submultiplicative norm such that $\|a^*\| = \|a\|$ for each $a \in \mathcal{A}$. If in addition $\mathcal{A}$ has a unit 1 such that $\|1\| = 1$, we call $\mathcal{A}$ a unital Banach $*$-algebra.

A C^*-algebra $\mathcal{A}$ is a Banach $*$-algebra which satisfies the following property:

$$\|aa^*\| = \|a\|^2 \quad \text{for all } a \in \mathcal{A}.$$

Definition 1.2.7. Let $\mathcal{H}$ be a complex vector space. A mapping $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is called an inner product in $\mathcal{H}$ if for any $x, y, z \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$ the following conditions are satisfied:
(i) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (the bar denotes the complex conjugate);
(ii) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$;
(iii) $\langle x, x \rangle \ge 0$;
(iv) $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.

A vector space with an inner product is called an **inner product space**.

Example 1.2.1. An important example of an inner product space is the space $\mathbb{C}^n$ of complex numbers. The inner product is defined by $\langle x, y \rangle = x\overline{y}$.

Theorem 1.2.1. *Let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be an inner product space. Define $\| \cdot \| : \mathcal{H} \to \mathbb{R}$ with $\|x\| = \sqrt{\langle x, x \rangle}$. Then*

(a) $|\langle x, y\rangle| \leq \|x\|\|y\|$ *for all* $x, y \in \mathcal{H}$;
(b) $(\mathcal{H}, \|\cdot\|)$ *is a normed space.*

Definition 1.2.8. A complete inner product space (with the norm induced by $\langle\cdot,\cdot\rangle$) is called a **Hilbert space**.

Definition 1.2.9. Let $\mathcal{X}$ be a vector space. A mapping $A : \mathcal{X} \to \mathcal{X}$ is said to be a **linear operator** if A satisfies the following for all $x, y \in \mathcal{X}$ and $\alpha \in \mathbb{C}$:
(i) $A(x + y) = Ax + Ay$;
(ii) $A(\alpha x) = \alpha Ax$.

Definition 1.2.10. Let $\mathcal{H}$ be a Hilbert space. A linear operator $A : \mathcal{H} \to \mathcal{H}$ is called bounded if there is a number $k > 0$ such that $\|Ax\| \leq k\|x\|$ for every $x \in \mathcal{H}$.

We denote the set of all bounded linear operators by $\mathbb{B}(\mathcal{H})$. If $A \in \mathbb{B}(\mathcal{H})$, then the usual **operator norm** is defined as

$$\|A\| = \sup\{\|Ax\| : x \in \mathcal{H}, \|x\| = 1\}.$$

Example 1.2.2. The simplest example of a bounded linear operator is the identity operator I.

Definition 1.2.11. Let $\mathcal{H}$ be a normed space. If a mapping $f : \mathcal{H} \to \mathbb{C}$ satisfies

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

for all $x, y \in \mathcal{H}$ and $\alpha \in \mathbb{C}$, then f is said to be a **linear functional.** A linear functional f is called bounded if there is a number $k > 0$ such that $|f(x)| \leq k\|x\|$ for every $x \in \mathcal{H}$. If f is a bounded linear functional, then

$$\|f\| = \sup\{|f(x)| : x \in \mathcal{H}, \|x\| = 1\}.$$

Theorem 1.2.2 (Riesz representation theorem). *Let f be a bounded linear functional on a Hilbert space $\mathcal{H}$. Then there exists a unique $x_0 \in \mathcal{H}$ such that $f(x) = \langle x, x_0\rangle$ for all $x \in \mathcal{H}$. Moreover, $\|f\| = \|x_0\|$.*

Definition 1.2.12. Let $A \in \mathbb{B}(\mathcal{H})$. For a fixed $y \in \mathcal{H}$, consider a bounded linear functional f such that $f(x) = \langle Ax, y\rangle$. By the Riesz representation theorem, there exists a unique $z \in \mathcal{H}$ such that $f(x) = \langle Ax, y\rangle = \langle x, z\rangle$ for each $x \in \mathcal{H}$. Then the operator $A^* : \mathcal{H} \to \mathcal{H}$ with $z = A^*y$ is a bounded linear operator on $\mathcal{H}$ and

$$\langle Ax, y\rangle = \langle x, z\rangle = \langle x, A^*y\rangle$$

for each $x \in \mathcal{H}$. We call A^* the adjoint of A.

Notice that $\mathbb{B}(\mathcal{H})$ with respect to involution $A \mapsto A^*$ and the usual operator norm is a C^*-algebra.

Definition 1.2.13. Let $A \in \mathbb{B}(\mathcal{H})$. Then:
(i) A is called an **invertible operator** if there exists an operator $B \in \mathbb{B}(\mathcal{H})$ such that $AB = BA = I$. This B is called inverse of A and is denoted by A^{-1}.
(ii) A is called a **self-adjoint operator** if $A = A^*$.
(iii) A is called a **normal operator** if it commutes with its adjoint, that is, $AA^* = A^*A$.
(iv) A is called a **unitary operator** if $AA^* = A^*A = I$.
(v) A is called a **positive operator** (we write $A \geq 0$) if it is self-adjoint and $\langle Ax, x\rangle \geq 0$ for all $x \in \mathcal{H}$.
(vi) A is called a **strictly positive operator** (we write $A > 0$) if A is an invertible operator and a positive operator.

Remark 1.2.1. For self-adjoint operators $A, B \in \mathbb{B}(\mathcal{H})$, $A \leq B$ means $B - A$ is positive.

Proposition 1.2.1. *Let $A \in \mathbb{B}(\mathcal{H})$.*
(a) *The operators $T_1 = A^*A$ and $T_2 = A + A^*$ are self-adjoint.*
(b) *The product of two self-adjoint operators is self-adjoint if and only if the operators commute.*

Theorem 1.2.3 (Cartesian form). *Let $T \in \mathbb{B}(\mathcal{H})$. Then there exist unique self-adjoint operators A and B such that $T = A + iB$ and $T^* = A - iB$.*

In the following, we collect some simple algebraic properties of invertible operators.

Proposition 1.2.2. *Let $A \in \mathbb{B}(\mathcal{H})$.*
(a) *The inverse of an invertible linear operator is a linear operator.*
(b) *An operator A is invertible if $Ax = 0$ implies $x = 0$.*
(c) *If operators A and B are invertible, then the operator AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*

Theorem 1.2.4. *If A is a normal operator, then $\|A^n\| = \|A\|^n$ for all $n \in \mathbb{N}$.*

Theorem 1.2.5. *Let $A \in \mathbb{B}(\mathcal{H})$ be self-adjoint. If $A \in \mathbb{B}(\mathcal{H})$ such that $A \geq cI$ for some $c > 0$, then A is invertible.*

Proposition 1.2.3. *Let $A \in \mathbb{B}(\mathcal{H})$.*
(a) *The operators A^*A and AA^* are positive.*
(b) *If A is a positive invertible operator, then its inverse A^{-1} is positive.*
(c) *The product of two commuting positive operators is a positive operator.*

Theorem 1.2.6. *Let $A, B, C \in \mathbb{B}(\mathcal{H})$.*
(a) *If $0 \leq A \leq B$, then $\|A\| \leq \|B\|$.*
(b) *If $0 < A \leq B$, then $B^{-1} \leq A^{-1}$.*
(c) *If $A \leq B$, then $C^*AC \leq C^*BC$.*

Theorem 1.2.7. *Let $A \in \mathbb{B}(\mathcal{H})$ be a positive operator. Then there exists a unique positive operator $B \in \mathbb{B}(\mathcal{H})$ such that $B^2 = A$ (denoted by $B = A^{\frac{1}{2}}$).*

Definition 1.2.14. The **spectrum** of a bounded operator A is the set

$$Sp(A) = \{\alpha \in \mathbb{C} : \ A - \alpha I \text{ is not invertible in } \mathbb{B}(\mathcal{H})\}.$$

Remark 1.2.2. Let $A \in \mathbb{B}(\mathcal{H})$ be a self-adjoint operator. Then, by the functional calculus, there exists a $*$-isometric isomorphism φ between the set $C(Sp(A))$ of all continuous functions on $Sp(A)$ and the C^*-algebra $C^*(A)$ generated by A and the identity operator I on $\mathcal{H}$ as follows: For $f, g \in C(Sp(A))$ and $\alpha, \beta \in \mathbb{C}$,
(a) $\varphi(\alpha f + \beta g) = \alpha\varphi(f) + \beta\varphi(g)$,
(b) $\varphi(fg) = \varphi(f)\varphi(g)$ and $\varphi(\bar{f}) = \varphi(f)^*$,
(c) $\|\varphi(f)\| = \|f\|$ where $\|f\| = \sup\limits_{t \in Sp(A)} |f(t)|$,
(d) $\varphi(f_0) = I$ and $\varphi(f_1) = A$ where $f_0(t) = 1$ and $f_1(t) = t$.

With this notation, we define $f(A) = \varphi(f)$ for all $f \in C(Sp(A))$.

Notice that, if $A \in \mathbb{B}(\mathcal{H})$ is a self-adjoint operator, then

$$f(t) \geq g(t), \quad t \in Sp(A) \Rightarrow f(A) \geq g(A) \tag{1.2.1}$$

provided that f and g are real-valued continuous functions.

Definition 1.2.15. A real-valued continuous function f on an interval J is said to be **operator monotone** if for all $A, B \in \mathbb{B}(\mathcal{H})$ with $Sp(A), Sp(B) \subseteq J$ such that $A \leq B$, the inequality $f(A) \leq f(B)$ holds. If $-f$ is operator monotone, then f is often called **operator monotone decreasing**.

An operator monotone function is said symmetric if $f(x) = xf(x^{-1})$, and normalized if $f(1) = 1$.

Definition 1.2.16. A real-valued continuous function f on an interval J is called **operator convex** (resp., **operator concave**) if $f((1-v)A + vB) \leq (1-v)f(A) + vf(B)$ (resp., $f((1-v)A + vB) \geq (1-v)f(A) + vf(B)$) holds for $0 \leq v \leq 1$ and self-adjoint operators $A, B \in \mathbb{B}(\mathcal{H})$ with $Sp(A), Sp(B) \subseteq J$.

Example 1.2.3. The function $f(t) = t^r$ is operator monotone on $(0, \infty)$ if and only if $0 \leq r \leq 1$. The function $f(t) = t^r$ is operator convex on $(0, \infty)$ if either $1 \leq r \leq 2$ or $-1 \leq r \leq 0$, and is operator concave on $(0, \infty)$ if $0 \leq r \leq 1$.

Example 1.2.4. The function $f(t) = \log t$ is operator monotone on $(0, \infty)$. Although, the function $f(t) = \exp t$ is neither operator convex nor operator monotone.

Remark 1.2.3. Let $f : (0, \infty) \to (0, \infty)$ be continuous. Then f is operator monotone if and only if f is operator concave.

Definition 1.2.17. A binary operation σ defined on the set of strictly positive operators is called an **operator mean** if:
(i) $I\sigma I = I$;
(ii) [transform inequality] $C^*(A\sigma B)C \le (C^*AC)\sigma(C^*BC)$;
(iii) [upper semicontinuity] $A_n \downarrow A$ and $B_n \downarrow B$ imply $(A_n\sigma B_n) \downarrow (A\sigma B)$, where $A_n \downarrow A$ means that $A_1 \ge A_2 \ge \cdots$ and $\|A_k x - Ax\| \to 0$ for any $x \in \mathcal{H}$ and a family $\{A_k\}$, $(k = 1, 2, \ldots)$ of $\mathbb{B}(\mathcal{H})$;
(iv) [monotonicity] $A \le B$ and $C \le D$ imply $A\sigma C \le B\sigma D$.

Let $f : J \subseteq \mathbb{R} \to \mathbb{R}$ be continuous and $A, B \in \mathbb{B}(\mathcal{H})$ be strictly positive operators such that $Sp(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}) \subseteq J$. Then the **operator connection** σ_f given by

$$A\sigma_f B = A^{\frac{1}{2}} f(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}},$$

is called f**-connection** (cf. [137]). An operator connection σ_f has properties such as monotonicity, transform inequality and upper semicontinuity. In addition, if an operator connection σ_f satisfies the condition $I\sigma_f I = I$, then σ_f is called an operator mean. f is often called a **representing function** of σ_f.

Definition 1.2.18. A linear map $\Phi : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ is said to be a **positive linear map** if $\Phi(A) \ge 0$ whenever $A \ge 0$ and Φ is called a **normalized linear map** (or **unital linear map**) if $\Phi(I_{\mathcal{H}}) = I_{\mathcal{K}}$ (we often use the symbol I for the **identity operator**).

In this book, we also treat some *matrix inequalities*. For such cases, we use the following notation. Let $M(n, \mathbb{C})$ be the set of all $n \times n$ matrices on the field $\mathbb{C}$. For a Hermitian matrix X, $X \ge 0$ means that $\langle\phi|X|\phi\rangle \ge 0$ for any vector $|\phi\rangle \in \mathbb{C}^n$. Where we used the so-called bra-ket notation which is often used in mathematical physics. The column vector (ket-vector) is represented by $|\phi\rangle \in \mathbb{C}^n$ and the row vector (bra-vector) is represented by $\langle\phi| = |\phi\rangle^t$ as transpose of column vector. So the norm of $|\phi\rangle$ is written as $\|\phi\| = \sqrt{\langle\phi|\phi\rangle}$ by an inner product $\langle\phi|\phi\rangle$.

Here, we denote the set of all **Hermitian matrices** by $M_h(n, \mathbb{C})$, that is, $M_h(n, \mathbb{C}) = \{X \in M(n, \mathbb{C}) | X = X^*\}$. We call X **skew-Hermitian** if $X^* = -X$.

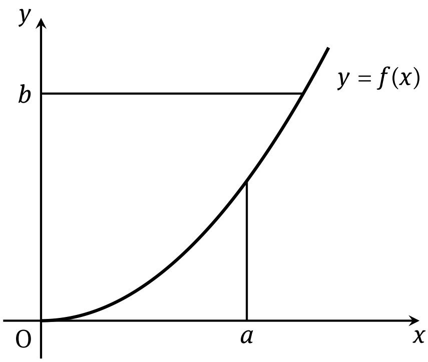
In addition, we denote the set of all **positive semidefinite matrices** by $M_+(n, \mathbb{C})$, that is, $M_+(n, \mathbb{C}) = \{X \in M_h(n, \mathbb{C}) | X \ge 0\}$. We also use the notation $X > 0$ if X is positive semi-definite and invertible matrix. A matrix $X > 0$ is called **positive definite matrix**. Let $M_{+,1}(n, \mathbb{C})$ be the set of strictly positive **density matrices**, that is, $M_{+,1}(n, \mathbb{C}) = \{\rho \in M_h(n, \mathbb{C}) \mid \mathrm{Tr}\,\rho = 1, \rho > 0\}$. If it is not otherwise specified in this book, we shall treat the case of faithful states, that is, $\rho > 0$.

2 Refinements and reverses for Young inequality

Let the continuous function $f : [0, \infty) \to [0, \infty)$ be monotone increasing with $f(0) = 0$. Then for $a, b > 0$, we have

$$ab \le \int_0^a f(x)\,dx + \int_0^b f^{-1}(x)\,dx$$

with equality if and only if $b = f(a)$. This inequality is often called **Young inequality**. The proof can be easily done by the ordering relation on the area in following figure:



If we take $f(x) = x^{p-1}$, $(p > 1)$ as an example, then its inverse is $f^{-1}(x) = x^{1/(p-1)}$. Thus we have

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}, \quad \left(p > 1, \frac{1}{p} + \frac{1}{q} = 1\right)$$

which implies $a^{1/p} b^{1/q} \le \frac{a}{p} + \frac{b}{q}$. Putting $v = 1/q$, we have the following first inequality:

$$(1-v)a + vb \ge a^{1-v} b^v \ge \left\{(1-v)a^{-1} + vb^{-1}\right\}^{-1}, \quad (0 \le v \le 1). \tag{2.0.1}$$

The second inequality is easily obtained by replacing a and b by $1/a$ and $1/b$, respectively. We often call the inequalities (2.0.1) the scalar (or numerical, classical) Young inequalities in the study of mathematical inequalities.

In addition, $a\nabla_v b = (1-v)a + vb$, $a\sharp_v b = a^{1-v} b^v$ and $a!_v b = \{(1-v)a^{-1} + vb^{-1}\}^{-1}$ are called the weighted **arithmetic mean**, **geometric mean** and **harmonic mean**, respectively.

There are other means such as logarithmic mean and Heron mean and so on. The **logarithmic mean** is defined by

$$L(a,b) = \frac{a-b}{\log a - \log b} \quad (a \ne b), \quad L(a,a) = a \tag{2.0.2}$$

https://doi.org/10.1515/9783110643473-002

for two positive real numbers a and b. The estimation on the Heron mean will be given in Section 3.2.

It is well known that we have the following operator (or matrix) Young inequalities for invertible positive operators A and B:

$$A\nabla_v B \geq A\sharp_v B \geq A!_v B, \tag{2.0.3}$$

where the weighted arithmetic mean, geometric mean and harmonic mean for positive operators A and B are defined for $v \in [0,1]$, respectively:

$$\begin{aligned} A\nabla_v B &= (1-v)A + vB, \\ A\sharp_v B &= A^{1/2}(A^{-1/2}BA^{-1/2})^v A^{1/2}, \\ A!_v B &= \{(1-v)A^{-1} + vB^{-1}\}^{-1}. \end{aligned} \tag{2.0.4}$$

We denote simply $A\nabla B = A\nabla_{\frac{1}{2}} B$, $A\sharp B = A\sharp_{\frac{1}{2}} B$ and $A!B = A\sharp_{\frac{1}{2}} B$. Throughout this book, we also use the symbol $A\natural_v B = A^{1/2}(A^{-1/2}BA^{-1/2})^v A^{1/2}$ for $v \in \mathbb{R}$ for wide range of $v \in [0,1]$.

We deal with the recent advances on refinements and reverses for the scalar/operator Young inequalities (2.0.1) and (2.0.3) in Chapter 2.

To obtain refinements and reverses for the Young inequalities, some constants will appear in the literature. The typical constants are the **Kantorovich constant** $K(h)$ and the **Specht ratio** $S(h)$. The Specht ratio $S(h)$ is defined by

$$S(h) = \frac{h^{\frac{1}{h-1}}}{e \log h^{\frac{1}{h-1}}}, \quad (h \neq 1) \tag{2.0.5}$$

for positive real number h. The Kantorovich constant is also defined by

$$K(h) = \frac{(h+1)^2}{4h} \tag{2.0.6}$$

for positive real number h. Furthermore, the **generalized Kantrovich constant** is often defined by

$$K(m,M,p) = \frac{(mM^p - Mm^p)}{(p-1)(M-m)}\left(\frac{p-1}{p}\frac{M^p - m^p}{mM^p - Mm^p}\right)^p, \quad 0 < m < M \tag{2.0.7}$$

or

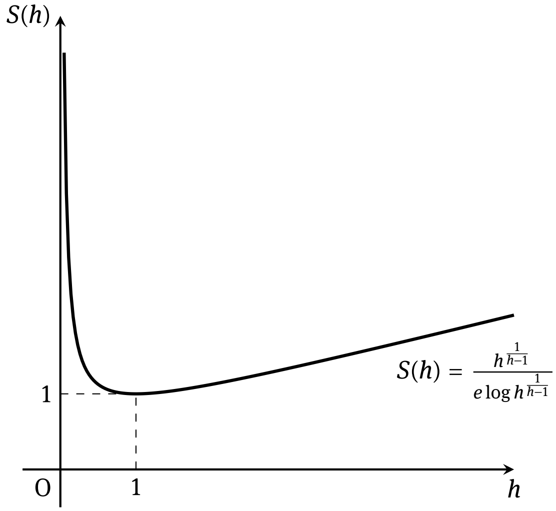
$$K(h,p) = \frac{(h^p - h)}{(p-1)(h-1)}\left(\frac{p-1}{p}\frac{h^p - 1}{h^p - h}\right)^p, \quad h > 0 \tag{2.0.8}$$

for any $p \in \mathbb{R}$ [94, Definition 1].

We here review the properties of the Specht ratio. See [226, 223, 51], for example, as for the proofs and the details.

Lemma 2.0.1. *The Specht ratio $S(h)$ has the following properties:*
(i) $S(1) = 1$ *and* $S(h) = S(1/h) > 1$ *for* $h > 0$ *and* $h \neq 1$.
(ii) $S(h)$ *is a monotone increasing function on* $(1, \infty)$.
(iii) $S(h)$ *is a monotone decreasing function on* $(0, 1)$.

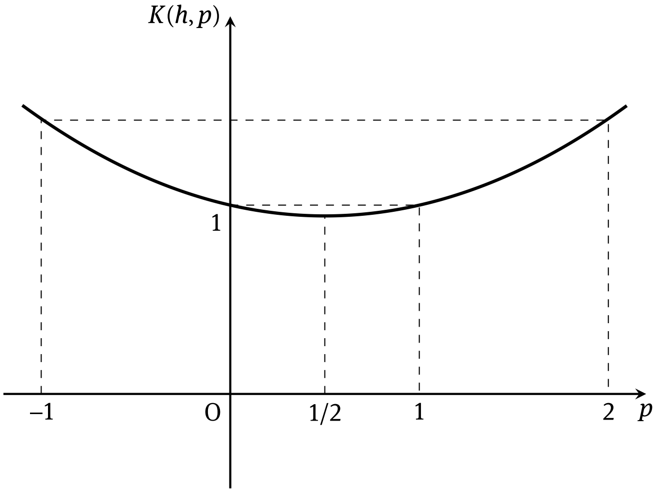
The Kantorovich constant $K(h)$ also has just same properties.



We also give properties of the generalized Kantorovich constant $K(h, p)$. See [97] for the proofs of Lemma 2.0.2

Lemma 2.0.2 ([97, Theorem 2.54]). *For $h > 0$ and $p \in \mathbb{R}$, the generalized Kantorovich constant $K(h, p)$ has the following properties:*
(i) $K(h, -1) = K(h, 2) = K(h)$, *where $K(h)$ is the original Kantorovich constant.*
(ii) *For any* $p > 0$, $K(1/h, p) = K(h, p)$.
(iii) *For any* $p > 0$, $K(h, 1 - p) = K(h, p)$. *(i. e., $K(h, p)$ is symmetry for $p = 1/2$.)*
(iv) $K(h, p)$ *is decreasing for $p < 1/2$ and increasing for $p > 1/2$. ($K(h, p)$ takes a minimum value $\frac{h^{1/4}}{h^{1/2}+1}$ when $p = 1/2$.)*
(v) $K(h, 0) = K(h, 1) = 1$.
(vi) *For any* $p \in \mathbb{R}$, $K(1, p) = 1$.
(vii) *For* $pr \neq 0$, $K(h^r, p/r)^{1/p} = K(h^p, r/p)^{-1/r}$.
(viii) *For any* $p \geq 1$ *and* $h > 1$, $K(h, p) \leq h^{p-1}$.

The following figure shows the graph of $K(h, p)$.



2.1 Reverses for Kittaneh–Manasrah inequality

The following inequalities are due to F. Kittaneh and Y. Manasrah. They are important and fundamental to start this topic. The proof is done for two cases, $0 \le v \le 1/2$ and $1/2 \le v \le 1$, with elementary calculations.

Proposition 2.1.1 ([130, 131]). *For $a, b \ge 0$ and $v \in [0,1]$, we have*

$$r(\sqrt{a} - \sqrt{b})^2 \le (1-v)a + vb - a^{1-v}b^v \le R(\sqrt{a} - \sqrt{b})^2, \tag{2.1.1}$$

where $r = \min\{v, 1-v\}$ and $R = \max\{v, 1-v\}$.

The inequalities given in Proposition 2.1.1 are often called the **Kittaneh–Manasrah inequality**. Note that this inequality has been given in [27]. Also similar but different type inequalities are known:

$$r^2(\sqrt{a} - \sqrt{b})^2 \le ((1-v)a + vb)^2 - (a^{1-v}b^v)^2 \le R^2(\sqrt{a} - \sqrt{b})^2,$$

where the first inequality above was given in [116] and the second one given in [131]. The first inequality in Proposition 2.1.1 clearly refines the first inequality of (2.0.1) for nonnegative real numbers a, b and $v \in [0,1]$.

It is not so difficult to give the operator Young inequalities based on Kittaneh–Manasrah inequality (2.1.1).

Proposition 2.1.2 ([64]). *For $v \in [0,1]$ and positive operators A and B, we have*

$$A\nabla_v B \ge A\sharp_v B + 2r(A\nabla B - A\sharp B) \ge A\sharp_v B \ge \left\{A^{-1}\sharp_v B^{-1} + 2r\left(A!B - A^{-1}\sharp B^{-1}\right)\right\}^{-1} \ge A!_v B$$

where $r = \min\{v, 1-v\}$.

As for the reverse inequalities of the Young inequality, M. Tominaga gave the following interesting operator inequalities. We call them **Tominaga's reverse Young inequalities**. He called them *converse* inequalities; however, we use the term *reverse* for such inequalities, throughout this book.

Proposition 2.1.3 ([226]). *Let $v \in [0,1]$, positive operators A and B such that $0 < mI \le A, B \le MI$ with $h = \frac{M}{m} > 1$. Then we have the following inequalities for every $v \in [0,1]$:*

(i) *(Reverse ratio inequality)*

$$S(h)A\sharp_v B \ge (1-v)A + vB, \tag{2.1.2}$$

where the constant $S(h)$ is called Specht ratio [223, 51] *and defined by* (2.0.5).

(ii) *(Reverse difference inequality)*

$$hL(1,h)\log S(h)B + A\sharp_v B \ge (1-v)A + vB, \tag{2.1.3}$$

where the logarithmic mean $L(\cdot,\cdot)$ is defined by (2.0.2).

In the following, we give reverse inequalities of the Kittaneh–Manasrah inequality for positive operators in both ratio and difference reverse inequalities, by the similar way to Tominaga's reverse Young inequalities.

For positive real numbers a, b and $v \in [0,1]$, M. Tominaga showed the following inequality [226]:

$$S\left(\frac{a}{b}\right)a^{1-v}b^{v} \ge (1-v)a + vb, \tag{2.1.4}$$

which is a ratio reverse inequality for the Young inequality in [226]. Combining the results in [130] and in [226], we show a **ratio reverse for the Kittaneh–Manasrah inequality** (2.1.1) in this section.

Lemma 2.1.1. *For $a, b > 0$ and $v \in [0,1]$, we have*

$$S\left(\sqrt{\frac{a}{b}}\right)a^{1-v}b^{v} \ge (1-v)a + vb - r(\sqrt{a} - \sqrt{b})^2, \tag{2.1.5}$$

where $r = \min\{v, 1-v\}$.

Applying Lemma 2.1.1, we have the ratio reverse of the refined Young inequality for positive operators.

Theorem 2.1.1 ([64]). *For $A, B > 0$ satisfy $0 < mI \le A, B \le MI$ with $m < M$ and for any $v \in [0,1]$, we then have*

$$S(\sqrt{h})A\sharp_v B \ge A\nabla_v B - 2r(A\nabla B - A\sharp B), \tag{2.1.6}$$

where $h = \frac{M}{m} > 1$ and $r = \min\{v, 1-v\}$.

Remark 2.1.1. We easily find that both sides in the inequality (2.1.5) is less than or equal to those in the inequality (2.1.4) so that neither the inequality (2.1.5) nor the inequality (2.1.4) is uniformly better than the other. In addition, we have no ordering between $S(\sqrt{\frac{a}{b}})a^{1-v}b^{v}$ and $(1-v)a+vb$.

For the classical Young inequality, the following reverse inequality is known. For $a, b > 0$ and $v \in [0,1]$, M. Tominaga showed the following inequality [226]:

$$L(a,b)\log S\left(\frac{a}{b}\right) \geq (1-v)a + vb - a^{1-v}b^{v}, \tag{2.1.7}$$

which is difference reverse for the Young inequality in [226]. Similarly, we can obtain the **difference reverse for Kittaneh–Manasrah inequality** given in (2.1.1).

Lemma 2.1.2. *For $a, b > 0$ and $v \in [0,1]$, we have*

$$\omega L(\sqrt{a}, \sqrt{b})\log S\left(\sqrt{\frac{a}{b}}\right) \geq (1-v)a + vb - a^{1-v}b^{v} - r(\sqrt{a}-\sqrt{b})^2, \tag{2.1.8}$$

where $\omega = \max\{\sqrt{a}, \sqrt{b}\}$.

We also have the following theorem. It can be proven by the similar method in [226] so that we omit the proof.

Theorem 2.1.2 ([64]). *For $A, B > 0$ satisfying $0 < mI \leq A, B \leq MI$ with $m < M$ and for any $v \in [0,1]$, we then have*

$$h\sqrt{M}L(\sqrt{M}, \sqrt{m})\log S(\sqrt{h}) \geq A\nabla_v B - A\sharp_v B - 2r(A\nabla B - A\sharp B), \tag{2.1.9}$$

where $h = \frac{M}{m} > 1$ and $r = \min\{v, 1-v\}$.

Remark 2.1.2. It is remarkable that we have no ordering between $L(a,b)\log S(\frac{a}{b})$ and $\omega L(\sqrt{a}, \sqrt{b})\log S(\sqrt{\frac{a}{b}}) + r(\sqrt{a}-\sqrt{b})^2$ for any $a, b > 0$ and $v \in [0,1]$. Therefore, we may claim that Theorem 2.1.2 is also nontrivial from the sense of finding a tighter upper bound of $(1-v)a + vb - a^{1-v}b^{v}$. See [64] for the details.

2.2 Young inequalities with Specht ratio

In section 2.1, we stated reverses for the Kittaneh–Manasrah inequality (2.1.1) due to F. Kittaneh and Y. Manasrah in [130] applying the similar method by M. Tominaga in [226]. Since the inequality (2.1.1) can be regarded as the difference-type refinement of Young inequality. In this section, we give a ratio-type refinement of Young inequality.

We use the following lemmas to show our theorem.

Lemma 2.2.1. *For $x \geq 1$, we have*

$$\frac{2(x-1)}{x+1} \leq \log x \leq \frac{x-1}{\sqrt{x}}. \tag{2.2.1}$$

Proof. It is known the following relation for three means:

$$\sqrt{xy} < \frac{x-y}{\log x - \log y} < \frac{x+y}{2}$$

for positive real numbers x and y, where $x \neq y$. This relation shows the lemma. □

Lemma 2.2.2. *For $t > 0$, we have*

$$e(t^2+1) \geq (t+1)t^{\frac{t}{t-1}}. \tag{2.2.2}$$

Proof. We put

$$f(t) = e(t^2+1) - (t+1)t^{\frac{t}{t-1}}.$$

By using the first inequality of (2.2.1) and by the fact that $\lim_{t\to 1} t^{\frac{1}{t-1}} = e$ and the function $t^{\frac{1}{t-1}}$ is monotone decreasing on $t \in [1,\infty)$, we have $f'(t) \geq 0$. (See [65] for details.) Thus with $f(1) = 0$ we have $f(t) \geq 0$ for $t \geq 1$. That is, we obtained the following inequality:

$$e(t^2+1) \geq (t+1)t^{\frac{t}{t-1}}, \quad t \geq 1.$$

Putting $t = \frac{1}{s}$ in the above inequality with simple calculations, we have

$$e(s^2+1) \geq (s+1)s^{\frac{s}{s-1}}, \quad 0 < s \leq 1.$$ □

Then we have the following inequality which improves the classical Young inequality. We call it a **refined Young inequality with Specht ratio**.

Theorem 2.2.1 ([65]). *For $a, b > 0$ and $v \in [0,1]$,*

$$(1-v)a + vb \geq S\left(\left(\frac{b}{a}\right)^r\right)a^{1-v}b^v, \tag{2.2.3}$$

where $r = \min\{v, 1-v\}$ and $S(\cdot)$ is the Specht ratio.

The proof can be done by the use of Lemma 2.2.1 and Lemma 2.2.2. We omit its proof. See [65] for it. The following inequality also improves the second inequality in (2.0.1).

Corollary 2.2.1. *For $a, b > 0$ and $v \in [0,1]$, we have*

$$a^{1-v}b^v \geq S\left(\left(\frac{a}{b}\right)^r\right)\left((1-v)a^{-1} + vb^{-1}\right)^{-1}, \tag{2.2.4}$$

where $r = \min\{v, 1-v\}$ and $S(\cdot)$ is the Specht ratio.

Proof. Replace a and b in Theorem 2.2.1 by $\frac{1}{a}$ and $\frac{1}{b}$, respectively. □

Applying Theorem 2.2.1, we have the following operator inequality for positive operators.

Theorem 2.2.2 ([65]). *For $A, B > 0$ and $m, m', M, M' > 0$ satisfying the following conditions either (i) or (ii):*
(i) $0 < m'I \leq A \leq mI < MI \leq B \leq M'I$;
(ii) $0 < m'I \leq B \leq mI < MI \leq A \leq M'I$;

with $h = \frac{M}{m}$, we have

$$A\nabla_v B \geq S(h^r)A\sharp_v B \geq A\sharp_v B \geq S(h^r)A!_v B \geq A!_v B$$

where $v \in [0,1]$, $r = \min\{v, 1-v\}$ and $S(\cdot)$ is the Specht ratio.

Theorem 2.2.2 corresponds to Proposition 2.1.2 as a ratio version. In the paper [64], we have proved the **refined Young inequality for n numbers**.

Proposition 2.2.1 ([64]). *Let $a_1, \ldots, a_n \geq 0$ and $p_1, \ldots, p_n > 0$ with $\sum_{j=1}^{n} p_j = 1$ and $\lambda = \min\{p_1, \ldots, p_n\}$. If we assume that the multiplicity attaining λ is 1, then we have*

$$\sum_{i=1}^{n} p_i a_i - \prod_{i=1}^{n} a_i^{p_i} \geq n\lambda\left(\frac{1}{n}\sum_{i=1}^{n} a_i - \prod_{i=1}^{n} a_i^{1/n}\right), \tag{2.2.5}$$

with equality if and only if $a_1 = \cdots = a_n$.

Proof. We suppose $\lambda = p_j$. For any $j = 1, \ldots, n$, we then have

$$\begin{aligned}\sum_{i=1}^{n} p_i a_i - p_j\left(\sum_{i=1}^{n} a_i - n\prod_{i=1}^{n} a_i^{1/n}\right) &= np_j\left(\prod_{i=1}^{n} a_i^{1/n}\right) + \sum_{i=1, i\neq j}^{n} (p_i - p_j)a_i \\ &\geq \prod_{i=1, i\neq j}^{n} (a_1^{1/n} \cdots a_n^{1/n})^{np_j} a_i^{p_i - p_j} \\ &= a_1^{p_1} \cdots a_n^{p_n}.\end{aligned}$$

In the above process, the arithmetic–geometric mean inequality was used. The equality holds if and only if

$$(a_1 a_2 \cdots a_n)^{\frac{1}{n}} = a_1 = a_2 = \cdots = a_{j-1} = a_{j+1} = \cdots = a_n$$

by the arithmetic–geometric mean inequality. Therefore, $a_1 = a_2 = \cdots = a_{j-1} = a_{j+1} = \cdots = a_n = a$, then we have $a_j^{\frac{1}{n}} a^{\frac{n-1}{n}} = a$ from the first equality. Thus we have $a_j = a$, which completes the proof. □

The inequality (2.2.5) is obviously improvement of the arithmetic–geometric mean inequality. To obtain the improved inequality (2.2.5), the arithmetic–geometric mean

inequality is used. Such an inequality is often called the **self-improvement inequality**. Such techniques will appear again in Section 4.1. It is also notable that we do not need the assumption that the multiplicity attaining λ is 1, to prove only inequality (2.2.5) without the equality condition. This assumption connects with the equality condition.

Closing this section, we give comments on the ratio refined Young inequality for n real numbers. We have not yet found its proof. We also have not found any counterexamples for the following 3-variables case:

$$w_1a_1 + w_2a_2 + w_3a_3 \geq S(h^r)a_1^{w_1}a_2^{w_2}a_3^{w_3}, \tag{2.2.6}$$

for $a_i \in [m, M]$ where $0 < m < M$ with $h = \frac{\max\{a_1,a_2,a_3\}}{\min\{a_1,a_2,a_3\}}$ and $r = \min\{w_1, w_2, w_3\}$, where $w_i > 0$ and $w_1 + w_2 + w_3 = 1$.

2.3 Reverse Young inequalities

Our purpose of this section is to give two *alternative* reverse inequalities which are different from Proposition 2.1.3. We first show the following remarkable scalar inequality.

Lemma 2.3.1 ([77, Theorem 2.1]). *Let $f : [a, b] \to \mathbb{R}$ be a twice differentiable function such that there exist real constant M such that $0 \leq f''(x) \leq M$ for $x \in [a, b]$.*

Then we have

$$0 \leq (1-v)f(a) + vf(b) - f\big((1-v)a + vb\big) \leq v(1-v)M(b-a)^2, \tag{2.3.1}$$

where $v \in [0, 1]$.

Proof. The first part of inequality (2.3.1) holds because f is a convex function. So, we prove second part of inequality (2.3.1). For $v \in \{0, 1\}$, we obtain the equality in relation (2.3.1). Now, we consider $v \in (0, 1)$, which means that $a < (1-v)a + vb < b$ and we use Lagrange's theorem for the intervals $[a, (1-v)a + vb]$ and $[(1-v)a + vb, b]$. Therefore, there exists real numbers $c_1 \in (a, (1-v)a + vb)$ and $c_2 \in ((1-v)a + vb, b)$ such that

$$f\big((1-v)a + vb\big) - f(a) = v(b-a)f'(c_1) \tag{2.3.2}$$

and

$$f(b) - f\big((1-v)a + vb\big) = (1-v)(b-a)f'(c_2). \tag{2.3.3}$$

Multiplying relation (2.3.2) by $(1-v)$ and relation (2.3.3) by v, and then adding, we deduce the following relation:

$$(1-v)f(a) + vf(b) - f\big((1-v)a + vb\big) = v(1-v)(b-a)\big(f'(c_2) - f'(c_1)\big).$$

Again, applying Lagrange's theorem on the interval $[c_1, c_2]$, we obtain

$$(1-v)f(a) + vf(b) - f\big((1-v)a + vb\big) = v(1-v)(b-a)(c_2 - c_1)f''(c), \tag{2.3.4}$$

where $c \in (c_1, c_2)$. Since $0 \le f''(x) \le M$ for $x \in [a, b]$ and $c_2 - c_1 \le b - a$ and making use of relation (2.3.4), we obtain the second inequality in (2.3.1). □

Corollary 2.3.1. *For $a, b > 0$ and $v \in [0, 1]$, the following inequalities hold:*
(i) *(Ratio-type reverse inequality)*

$$a^{1-v}b^v \le (1-v)a + vb \le a^{1-v}b^v \exp\left(\frac{v(1-v)(a-b)^2}{d_1^2}\right), \tag{2.3.5}$$

where $d_1 = \min\{a, b\}$.
(ii) *(Difference-type reverse inequality)*

$$a^{1-v}b^v \le (1-v)a + vb \le a^{1-v}b^v + v(1-v)\left(\log\left(\frac{a}{b}\right)\right)^2 d_2, \tag{2.3.6}$$

where $d_2 = \max\{a, b\}$.

Proof.
(i) It is easy to see that if we take $f(x) = -\log x$ in Theorem 2.3.1, then we have

$$\log\{(1-v)a + vb\} \le \log(a^{1-v}b^v) + \log\big(\exp\{v(1-v)f''(c)(b-a)^2\}\big)$$

which implies inequality (2.3.5), since $f''(c) = \frac{1}{c^2} \le \frac{1}{d_1^2}$.
(ii) If we take $f(x) = e^x$ (which is convex on $(-\infty, \infty)$) in Theorem 2.3.1, we obtain the relation

$$0 \le ve^\alpha + (1-v)e^\beta - e^{v\alpha + (1-v)\beta} \le v(1-v)(\alpha - \beta)^2 f''(\gamma),$$

where $\gamma = \max\{\alpha, \beta\}$. Putting $a = e^\alpha$ and $b = e^\beta$, then we have

$$0 \le (1-v)a + vb - a^{1-v}b^v \le v(1-v)e^c\left(\log\frac{a}{b}\right)^2,$$

where $c = \max\{\log a, \log b\}$. Thus we have inequality (2.3.6), putting $d_2 = e^c$. □

Both second inequalities are different from Proposition 2.1.3. And they represent alternative reverses for Young inequality. We can obtain the following operator inequalities using Corollary 2.3.1.

Theorem 2.3.1 ([77]). *For $v \in [0, 1]$, two invertible positive operators A and B satisfying the ordering $mI \le A \le B \le MI \le I$ with $h = \frac{M}{m}$, we have the following operator inequalities:*

(i) *(Ratio-type reverse inequality)*

$$A\sharp_v B \le (1-v)A + vB \le \exp\big(v(1-v)(1-h)^2\big)A\sharp_v B. \tag{2.3.7}$$

(ii) *(Difference-type reverse inequality)*

$$A\sharp_v B \le (1-v)A + vB \le A\sharp_v B + v(1-v)(\log h)^2 B. \tag{2.3.8}$$

Remark 2.3.1. It is natural to consider that our inequalities are better than Tominaga's inequalities given in Proposition 2.1.3, under the assumption $A \le B$. However, there is no ordering between Theorem 2.3.1 and Proposition 2.1.3, by some numerical computations [77]. Therefore, we may conclude two reverse inequalities for Young's inequality given in Theorem 2.3.1 are not trivial results under the assumption $A \le B$.

2.4 Generalized reverse Young inequalities

After the appearance of the Kittaneh–Manasrah inequality, some refinements have been done for Young inequality, by dividing the interval of the weighted parameter v and recently M. Sababheh and D. Choi gave its complete refinement [213]. As for the reverses of Young inequality for the divided interval, the results given by S. Furuichi, M. B. Ghaemi and N. Gharakhanlu in [73] are known. In this section, we give *alternative proofs* for the generalized reverse Young inequalities shown in [73] by the method given in [69].

Theorem 2.4.1 ([73, Theorem 1], [69, Theorem 1.1]). *Let v, t be real numbers with $t > 0$, and $n \in \mathbb{N}$ with $n \ge 2$.*

(i) *For $v \notin [\frac{1}{2}, \frac{2^{n-1}+1}{2^n}]$, we have*

$$(1-v) + vt \le t^v + (1-v)(1-\sqrt{t})^2 + (2v-1)\sqrt{t}\sum_{k=2}^{n} 2^{k-2}\big(\sqrt[2^k]{t}-1\big)^2. \tag{2.4.1}$$

(ii) *For $v \notin [\frac{2^{n-1}-1}{2^n}, \frac{1}{2}]$, we have*

$$(1-v)t + v \le t^{1-v} + v(\sqrt{t}-1)^2 + (1-2v)\sqrt{t}\sum_{k=2}^{n} 2^{k-2}\big(\sqrt[2^k]{t}-1\big)^2. \tag{2.4.2}$$

Proof.

(i) Direct calculations imply

$$\begin{aligned}(1-v) + vt &- (1-v)(1-\sqrt{t})^2 - (2v-1)\sqrt{t}\sum_{k=2}^{n} 2^{k-2}\big(\sqrt[2^k]{t}-1\big)^2 \\ &= \sqrt{t} + \sqrt{t}\left(v - \frac{1}{2}\right)2^n\big(\sqrt[2^n]{t}-1\big).\end{aligned} \tag{2.4.3}$$

Thus the inequality (2.4.1) is equivalent to the inequality

$$\left(v-\frac{1}{2}\right)2^n(\sqrt[2^n]{t}-1) \le t^{v-\frac{1}{2}}-1. \tag{2.4.4}$$

This inequality is true by (i) of Lemma 2.4.2 below with $r=\frac{1}{2^n}$, since the conditions $0<r<v-\frac{1}{2}$ or $v-\frac{1}{2}<0<r$ are satisfied in the case of $v \notin [\frac{1}{2}, \frac{2^{n-1}+1}{2^n}]$ in (i) of Lemma 2.4.2.

(ii) Exchanging $1-v$ with v in (i) of Lemma 2.4.2, we have

$$\left(\frac{1}{2}-v\right)\frac{t^r-1}{r} \le t^{\frac{1}{2}-v}-1 \tag{2.4.5}$$

for $v<\frac{1}{2}-r$ or $\frac{1}{2}<v$. Exchanging $1-v$ with v in the inequality (2.4.4), we have

$$\left(\frac{1}{2}-v\right)2^n(\sqrt[2^n]{t}-1) \le t^{\frac{1}{2}-v}-1. \tag{2.4.6}$$

This inequality is true by the inequality (2.4.5) with $r=\frac{1}{2^n}$, since the conditions $v<\frac{1}{2}-r$ or $\frac{1}{2}<v$ are satisfied in the case of $v \notin [\frac{2^{n-1}-1}{2^n}, \frac{1}{2}]$ in (i) of Lemma 2.4.2. □

Theorem 2.4.2 ([73, Theorem 3], [69, Theorem 1.2]). *Let v, t be real numbers with $t>0$, and $n \in \mathbb{N}$.*

(i) *For $v \notin [0, \frac{1}{2^n}]$, we have*

$$(1-v)+vt \le t^v + v\sum_{k=1}^{n} 2^{k-1}(1-\sqrt[2^k]{t})^2. \tag{2.4.7}$$

(ii) *For $v \notin [\frac{2^n-1}{2^n}, 1]$, we have*

$$(1-v)t+v \le t^{1-v} + (1-v)\sum_{k=1}^{n} 2^{k-1}(1-\sqrt[2^k]{t})^2. \tag{2.4.8}$$

Proof.

(i) Direct calculations imply

$$(1-v)+vt-v\sum_{k=1}^{n} 2^{k-1}(1-\sqrt[2^k]{t})^2 = v2^n(\sqrt[2^n]{t}-1)+1 \tag{2.4.9}$$

so that the inequality (2.4.7) is equivalent to the inequality

$$v2^n(\sqrt[2^n]{t}-1) \le t^v - 1. \tag{2.4.10}$$

This inequality is true by (ii) of Lemma 2.4.2 with $r=\frac{1}{2^n}$, since the conditions $0<r<v$ or $v<0<r$ are satisfied in the case of $v \notin [0, \frac{1}{2^n}]$ in (ii) of Lemma 2.4.2.

(ii) Exchanging $1-v$ with v in (ii) of Lemma 2.4.2, we have

$$(1-v)\frac{t^r-1}{r} \le t^{1-v}-1 \tag{2.4.11}$$

for $0<r<1-v$ or $1-v<0<r$. Exchanging $1-v$ with v in the inequality (2.4.10), we also have

$$(1-v)2^n(\sqrt[2^n]{t}-1) \le t^{1-v}-1. \tag{2.4.12}$$

This inequality is true by the inequality (2.4.11) with $r=\frac{1}{2^n}$, since the conditions $0<r<1-v$ or $1-v<0<r$ are satisfied in the case of $v \notin [\frac{2^n-1}{2^n},1]$ in (ii) of Lemma 2.4.2. □

To complete our elementary proofs for the above theorems, we need Lemma 2.4.2. So we start from the famous formula

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \lim_{r\to 0}(1+rx)^{1/r}.$$

In this book, we often consider the inverse function of r-**exponential function** $\exp_r(x) = (1+rx)^{1/r}$, namely r-**logarithmic function** defined by $\ln_r x = \frac{x^r-1}{r}$ for $x>0$ and a real number $r \ne 0$.

Lemma 2.4.1. *$\ln_r x$ is a monotone increasing function in r.*

Proof. In the inequality $\log t \le t-1$ for $t>0$, we set $t=x^{-r}$ and we obtain the following:

$$\frac{\partial \ln_r x}{\partial r} = \frac{x^r(\log x^r - 1 + x^{-r})}{r^2} \ge 0.$$

□

Lemma 2.4.1 implies the following lemma.

Lemma 2.4.2. *Let r,v,t be real numbers with $r \ne 0$ and $t>0$.*

(i) *For $0<r<v-\frac{1}{2}$ or $v-\frac{1}{2}<0<r$, we have*

$$\left(v-\frac{1}{2}\right)\frac{t^r-1}{r} \le t^{v-\frac{1}{2}}-1. \tag{2.4.13}$$

(ii) *For $0<r<v$ or $v<0<r$, we have*

$$v\frac{t^r-1}{r} \le t^v-1. \tag{2.4.14}$$

Lemma 2.4.2 was a key tool to give alternative simple proofs for Theorem 2.4.1 ([73, Theorem 1]) and Theorem 2.4.2 ([73, Theorem 3]) without using complicated computations and the supplemental Young's inequality given in [73, Lemma 5] which used to prove [73, Theorem 1] and [73, Theorem 3].

By theory of Kubo–Ando, we have the following corollary from Lemma 2.4.2.

Corollary 2.4.1. *Let r, v be real numbers with $r \neq 0$. For $A, B > 0$, we have the following operator inequalities:*
(i) *For $0 < r < v - \frac{1}{2}$ or $v - \frac{1}{2} < 0 < r$, we have*

$$\left(v - \frac{1}{2}\right)\frac{A \natural_r B - A}{r} \leq A \natural_{v-\frac{1}{2}} B - A. \tag{2.4.15}$$

(ii) *For $0 < r < v$ or $v < 0 < r$, we have*

$$v\frac{A \natural_r B - A}{r} \leq A \natural_v B - A. \tag{2.4.16}$$

The methods in the above are applicable to obtain the inequalities in the following propositions, as additional results. They give refinements for Young inequality with divided infinitesimal intervals for sufficiently large n. The proofs are done by the use of Lemma 2.4.1 similar to Theorem 2.4.1 and 2.4.2 so that we omit their proofs. See [69] for the proofs. The following propositions are covered by the results in [213]. However as our advantage, our proofs given in [69] similar to the proofs for Theorem 2.4.1 and 2.4.2 are quite simple.

Proposition 2.4.1. *Let v, t be real numbers with $t > 0$ and $n \in \mathbb{N}$.*
(i) *For $v \in [0, \frac{1}{2^n}]$, we have*

$$(1 - v) + vt \geq t^v + v \sum_{k=1}^{n} 2^{k-1}(1 - \sqrt[2^k]{t})^2. \tag{2.4.17}$$

(ii) *For $v \in [\frac{2^n-1}{2^n}, 1]$, we have*

$$(1 - v)t + v \geq t^{1-v} + (1 - v) \sum_{k=1}^{n} 2^{k-1}(1 - \sqrt[2^k]{t})^2. \tag{2.4.18}$$

Proposition 2.4.2. *Let v, t be real numbers with $t > 0$, and $n \in \mathbb{N}$ with $n \geq 2$.*
(i) *For $v \in [\frac{1}{2}, \frac{2^{n-1}+1}{2^n}]$, we have*

$$(1 - v) + vt \geq t^v + (1 - v)(1 - \sqrt{t})^2 + (2v - 1)\sqrt{t} \sum_{k=2}^{n} 2^{k-2}(\sqrt[2^k]{t} - 1)^2. \tag{2.4.19}$$

(ii) *For $v \in [\frac{2^{n-1}-1}{2^n}, \frac{1}{2}]$, we have*

$$(1 - v)t + v \geq t^{1-v} + v(\sqrt{t} - 1)^2 + (1 - 2v)\sqrt{t} \sum_{k=2}^{n} 2^{k-2}(\sqrt[2^k]{t} - 1)^2. \tag{2.4.20}$$

The operator versions for Theorem 2.4.1, Theorem 2.4.2, Proposition 2.4.1 and Proposition 2.4.2 can be also easily shown by Kubo–Ando theory (or the standard functional calculus) in [73]. We omit them here.

2.5 Young inequalities by exponential functions

We study refined and reverse Young inequality with exponential functions in this section. As we have seen, a refined Young inequality with Specht ratio was shown in [65]:

$$\frac{(1-\nu)+\nu t}{t^{\nu}} \geq S(t^{r}) \tag{2.5.1}$$

for $t > 0$, where $r = \min\{\nu, 1-\nu\}$ with $\nu \in [0,1]$. The inequality (2.5.1) is a refinement of Young inequality in the sense of $S(h) \geq 1$ for $h > 0$.

On the other hand, the reverse Young inequality was given in [223, 226] for $t > 0$:

$$\frac{(1-\nu)+\nu t}{t^{\nu}} \leq S(t). \tag{2.5.2}$$

Therefore, as a quite natural insight, the following first inequality was conjectured in [43, 42] for $t > 0$ (the second inequality is trivial):

$$\frac{(1-\nu)+\nu t}{t^{\nu}} \leq S(t^{R}) \leq S(t), \tag{2.5.3}$$

where $R = \max\{\nu, 1-\nu\}$ with $\nu \in [0,1]$.

However, we have counterexamples for the first inequality in (2.5.3). Actually, we set $t = 2$ and $\nu = \frac{1}{2}$ for simply, then the inequality (2.5.3) becomes

$$\frac{3}{2} \leq S(\sqrt{2})\sqrt{2}.$$

By the numerical computations $S(\sqrt{2})\sqrt{2} \approx 1.43557$ so that the first inequality in (2.5.3) does not hold in general. (For supplementation, $S(2)\sqrt{2} \approx 1.50115$.) The above result was reported in [70].

2.5.1 Elementary proofs for Dragomir's inequalities

S. S. Dragomir established the following refinement of Young inequality in [43]. Here, we give an alternative and elementary proof for the following **Dragomir's reverse Young inequality** with exponential function.

Theorem 2.5.1 ([43], [71, Theorem 2.1]). *For $t > 0$ and $\nu \in [0,1]$,*

$$\frac{(1-\nu)+\nu t}{t^{\nu}} \leq \exp\left(\nu(1-\nu)\frac{(t-1)^2}{t}\right). \tag{2.5.4}$$

Proof. To prove the inequality (2.5.4), we put

$$f_{\nu}(t) = (1-\nu)t^{-\nu} + \nu t^{1-\nu} - \exp\left(\nu(1-\nu)\frac{(t-1)^2}{t}\right).$$

Then we calculate

$$\frac{df_v(t)}{dt} = v(1-v)(1-t)t^{-v-1}h_v(t),$$

where

$$h_v(t) = t^{v-1}(t+1)\exp\left(v(1-v)\frac{(t-1)^2}{t}\right) - 1 \geq 0.$$

The last inequality is due to Lemma 2.5.1 in the below. Therefore, we have $\frac{df_v(t)}{dt} \geq 0$ if $0 < t \leq 1$ and $\frac{df_v(t)}{dt} \leq 0$ if $t \geq 1$. Thus we have $f_v(t) \leq f_v(1) = 0$ which implies the inequality (2.5.4). □

Lemma 2.5.1. *For $t > 0$ and $v \in [0,1]$, we have*

$$t^{v-1}(t+1)\exp\left(v(1-v)\frac{(t-1)^2}{t}\right) \geq 1.$$

Proof. For any $t > 0$ and $0 \leq v \leq 1$, $\exp(v(1-v)\frac{(t-1)^2}{t}) \geq 1$. In addition, $t^{v-1}(t+1) = t^v + t^{v-1} \geq 1$ since $t^v \geq 1$, $t^{v-1} > 0$ for $t \geq 1$ and $t^v > 0$, $t^{v-1} \geq 1$ for $0 < t \leq 1$. Therefore, we have the desired inequality. □

Remark 2.5.1. It is known the following inequalities (see [248, 140]),

$$K^r(t) \leq \frac{(1-v)+vt}{t^v} \leq K^R(t), \tag{2.5.5}$$

where $r = \min\{v, 1-v\}$ and $R = \max\{v, 1-v\}$. The first inequality of (2.5.5) was shown by H. Zuo, G. Shi and M. Fujii. The second inequality of (2.5.5) was shown by W. Liao, J. Wu and J. Zhao. We call (2.5.5) the **Zuo–Liao inequality** for the convenience of the readers. By the numerical computations, we find that there is no ordering between $\exp(v(1-v)\frac{(t-1)^2}{t})$ and $K^R(t)$ so that Theorem 2.5.1 is not trivial one. See [71] for the details.

S. S. Dragomir also established the following refined Young inequalities with the general inequalities in his paper [42]. The following is our alternative and elementary proof for the **Dragomir's refined and reverse Young inequality** with exponential function.

Theorem 2.5.2 ([42], [71, Theorem 2.4]). *Let $t > 0$ and $v \in [0,1]$.*

(i) *If $0 < t \leq 1$, then*

$$\exp\left(\frac{v(1-v)}{2}(t-1)^2\right) \leq \frac{(1-v)+vt}{t^v} \leq \exp\left(\frac{v(1-v)}{2}\left(\frac{1}{t}-1\right)^2\right). \tag{2.5.6}$$

(ii) *If $t \geq 1$, then*

$$\exp\left(\frac{v(1-v)}{2}\left(\frac{1}{t}-1\right)^2\right) \leq \frac{(1-v)+vt}{t^v} \leq \exp\left(\frac{v(1-v)}{2}(t-1)^2\right). \tag{2.5.7}$$

Proof.

(i) We prove the first inequality of (2.5.6). To do this, we set

$$f_1(t,v) = (1-v)t^{-v} + vt^{1-v} - \exp\left(\frac{v(1-v)(t-1)^2}{2}\right).$$

Then we calculate

$$\frac{df_1(t,v)}{dt} = v(1-v)(1-t)h_1(t,v),$$

where

$$h_1(v,t) = \exp\left(\frac{v(1-v)(t-1)^2}{2}\right) - t^{-v-1}.$$

From Lemma 2.5.2 in the below, $h_1(t,v) \le 0$ which means $\frac{df_1(t,v)}{dt} \le 0$. Thus we have $f_1(t,v) \ge f_1(1,v) = 0$ which means the first inequality of (2.5.6) hold for $0 < t \le 1$ and $0 \le v \le 1$.

We prove the second inequality of (2.5.6). To do this, we set

$$f_2(t,v) = (1-v)t^{-v} + vt^{1-v} - \exp\left(\frac{v(1-v)(t-1)^2}{2t^2}\right).$$

Then we calculate

$$\frac{df_2(t,v)}{dt} = v(1-v)(1-t)t^{-3}h_2(t,v),$$

where

$$h_2(v,t) = \exp\left(\frac{v(1-v)(t-1)^2}{2t^2}\right) - t^{2-v}.$$

Since $\exp(u) \ge 1+u$ for $u \ge 0$, we have $h_2(v,t) \ge g_2(v,t)$, where

$$g_2(t,v) = 1 + \frac{v(1-v)(t-1)^2}{2t^2} - t^{2-v}.$$

Then we calculate

$$\frac{dg_2(t,v)}{dt} = t^{-3}\{(v-2)t^{4-v} + v(1-v)(t-1)\} \le 0$$

for $0 < t \le 1$ and $0 \le v \le 1$. Thus we have $g_2(t,v) \ge g_2(1,v) = 0$ which implies $h_2(t,v) \ge 0$ which means $\frac{df_2(t,v)}{dt} \ge 0$. Thus we have $f_2(t,v) \le f_2(1,v) = 0$ which means the second inequality of (2.5.6) hold for $0 < t \le 1$ and $0 \le v \le 1$.

(ii) We prove the double inequalities (2.5.7). Put $t = 1/s$ for $0 < t \le 1$ in (i). Then we have

$$\exp\left(\frac{v(1-v)}{2}\left(\frac{1}{s}-1\right)^2\right) \le \frac{(1-v)+v/s}{s^{-v}} \le \exp\left(\frac{v(1-v)}{2}(s-1)^2\right), \quad (s \ge 1)$$

which implies (2.5.7) by replacing v with $1-v$. □

Lemma 2.5.2. *For* $0 < t \le 1$ *and* $0 \le v \le 1$, *we have*

$$t^{v+1} \exp\left(\frac{v(1-v)}{2}(t-1)^2\right) \le 1. \tag{2.5.8}$$

Proof. We set the function as

$$f_v(t) = (v+1)\log t + \frac{v(1-v)}{2}(t-1)^2.$$

Then we calculate

$$\frac{df_v(t)}{dt} = \frac{v+1}{t} + v(1-v)(t-1), \quad \frac{d^2 f(t,v)}{dt^2} = -\frac{v+1}{t^2} + v(1-v), \quad \frac{d^3 f(t,v)}{dt^3} = \frac{2(v+1)}{t^3} \ge 0.$$

Thus we have $\frac{d^2 f_v(t)}{dt^2} \le \frac{d^2 f_v(1)}{dt^2} = -v^2 - 1 \le 0$ so that we have $\frac{df_v(t)}{dt} \ge \frac{df_v(1)}{dt} = v + 1 \ge 0$. Therefore, we have $f_v(t) \le f_v(1) = 0$ which implies the inequality (2.5.8). □

We obtained the following results in the paper [79]. See also Theorem 2.6.1 in Section 2.6.

Proposition 2.5.1 ([79]). *For* $0 \le v \le 1$ *and* $0 < t \le 1$, *we have*

$$m_v(t) \le \frac{(1-v) + vt}{t^v} \le M_v(t),$$

where

$$m_v(t) = 1 + \frac{v(1-v)(t-1)^2}{2}\left(\frac{t+1}{2}\right)^{-v-1}, \quad M_v(t) = 1 + \frac{v(1-v)(t-1)^2}{2} t^{-v-1}.$$

Remark 2.5.2. As shown in the paper [79], we have the inequality

$$M_v(t) \le \exp\left(v(1-v)\frac{(t-1)^2}{t}\right)$$

for $0 < t \le 1$ and $0 \le v \le \frac{1}{2}$. That is, the second inequality in Proposition 2.5.1 gives better bound than the inequality (2.5.4), in case of $0 < t \le 1$ and $0 \le v \le \frac{1}{2}$.

In the following proposition, we give the comparison on bounds in (i) of Theorem 2.5.2 and in Proposition 2.5.1.

Proposition 2.5.2. *For* $0 \le v \le 1$ *and* $0 < t \le 1$, *we have*

$$M_v(t) \le \exp\left(\frac{v(1-v)}{2}\left(\frac{1}{t} - 1\right)^2\right) \tag{2.5.9}$$

and

$$\exp\left(\frac{v(1-v)}{2}(t-1)^2\right) \le m_v(t). \tag{2.5.10}$$

Proof. We use the inequality

$$\exp(x) \geq 1 + x + \frac{1}{2}x^2, \quad (x \geq 0).$$

Then we calculate

$$\begin{aligned}
&\exp\left(\frac{v(1-v)}{2}\left(\frac{1}{t}-1\right)^2\right) - 1 - \frac{v(1-v)(t-1)^2}{2t^{v+1}} \\
&\geq \frac{v(1-v)(t-1)^2}{2t^2}\left(\frac{t^{v-1}-1}{t^{v-1}} + \frac{v(1-v)(t-1)^2}{4t^2}\right) \geq 0
\end{aligned}$$

for $0 \leq v \leq 1$ and $0 < t \leq 1$. Thus the inequality (2.5.9) was proved.

Putting $s = \frac{t+1}{2}$, the inequality (2.5.10) is equivalent to the inequality

$$\exp\left(2v(1-v)(s-1)^2\right) \leq 1 + 2v(1-v)(s-1)^2 s^{-v-1}, \quad \left(\frac{1}{2} < s \leq 1, 0 \leq v \leq 1\right). \tag{2.5.11}$$

For the special case $v = 0, 1$ or $s = 1$, the equality holds in (2.5.11) so that we assume $2v(1-v)(s-1)^2 \neq 0$. Then we use the inequality

$$\exp(x) < \frac{1}{1-x}, \quad (x < 1).$$

We calculate

$$\begin{aligned}
m_v(t) - \exp\left(\frac{v(1-v)}{2}(t-1)^2\right) &= 1 + 2v(1-v)(s-1)^2 s^{-v-1} - \exp\left(2v(1-v)(s-1)^2\right) \\
&> 1 + 2v(1-v)(s-1)^2 s^{-v-1} - \frac{1}{1-2v(1-v)(s-1)^2} \\
&= \frac{2v(1-v)(s-1)^2 s^{-v-1} g_v(s)}{1-2v(1-v)(s-1)^2},
\end{aligned}$$

where

$$g_v(s) = 1 - s^{v+1} - 2v(1-v)(s-1)^2.$$

We prove $g_v(s) \geq 0$. To this end, we calculate

$$\begin{aligned}
&g_v'(s) = -(v+1)s^v - 4v(1-v)(s-1), \quad g_v''(s) = -v(v+1)s^{v-1} - 4v(1-v), \\
&g_v^{(3)}(s) = v(1-v)(v+1)s^{v-2} \geq 0.
\end{aligned}$$

Thus we have $g_v''(s) \leq g_v''(1) = v(3v-5) \leq 0$ so that we have $g_v'(s) \leq g_v'(1/2) = 2h(v)$, where $h(v) = v(1-v) - \frac{v+1}{2^{v+1}}$. From Lemma 2.5.3, $h(v) \leq 0$ for $0 \leq v \leq 1$. Thus we have $g_v'(s) \leq 0$ which implies $g_v(s) \geq g_v(1) = 0$. Therefore, we have

$$m_v(t) - \exp\left(\frac{v(1-v)}{2}(t-1)^2\right) > 0$$

for $0 < v < 1$ and $0 < t < 1$. Taking account for the equality cases happen if $v = 0, 1$ or $t = 1$, we have the inequality (2.5.10). □

Lemma 2.5.3. *For* $0 \le v \le 1$, *we have* $\frac{v+1}{2^{v+1}} \ge v(1-v)$.

Proof. Since $v(1-v) \le \frac{1}{4}$ for $0 \le v \le 1$, it is sufficient to prove $\frac{v+1}{2^{v+1}} \ge \frac{1}{4}$. So we put $l(v) = 2(v+1) - 2^v$. Then we have $l''(v) = -(\log 2)^2 2^v \le 0$, $l(0) = 1$ and $l(1) = 2$. Therefore, we have $l(v) \ge 0$. □

Remark 2.5.3. As for the bounds on the ratio of arithmetic mean to geometric mean $\frac{(1-v)+vt}{t^v}$, Proposition 2.5.2 shows Proposition 2.5.1 is better than (i) of Theorem 2.5.2, for the case $0 < t \le 1$.

2.5.2 Further refinements by *r*-exponential function

We give a new refinement of Young inequality which is a further improvement of Theorem 2.5.1. Throughout this book, we use the one-parameter extended exponential function defined by $\exp_r(x) = (1 + rx)^{1/r}$ for $x > 0$ and $-1 \le r \le 1$ with $r \ne 0$ under the assumption that $1 + rx \ge 0$. We call $\exp_r(\cdot)$ ***r*-exponential function.**

Theorem 2.5.3 ([71, Theorem 3.1]). *For* $t > 0$ *and* $v \in [0,1]$,

$$\frac{(1-v)+vt}{t^v} \le 1 + v(1-v)\frac{(t-1)^2}{t}. \tag{2.5.12}$$

Proof. We set the function

$$f_v(t) = 1 + v(1-v)\frac{(t-1)^2}{t} - \frac{(1-v)+vt}{t^v}$$

for $t > 0$ and $0 \le v \le 1$. Then we have

$$\frac{df_v(t)}{dt} = \frac{v(1-v)(t-1)}{t^{v+1}}(t^v + t^{v-1} - 1).$$

Since $t^v \ge 1$ for $t \ge 1$ and $t^{v-1} \ge 1$ for $0 < t \le 1$, $t^v + t^{v-1} - 1 \ge 0$. Thus we have $\frac{df_v(t)}{dt} = 0$ when $t = 1$ and $\frac{df_v(t)}{dt} \le 0$ for $0 < t < 1$, and $\frac{df_v(t)}{dt} \ge 0$ for $t > 1$. Therefore, we have $f_v(t) \ge f_v(1) = 0$. □

Lemma 2.5.4. *The function* $\exp_r(x)$ *defined for* $x > 0$ *and* $r \in (0,1]$, *is monotone decreasing in* r.

Proof. We calculate $\frac{d\exp_r(x)}{dx} = \frac{(1+rx)^{\frac{1-r}{r}} g(rx)}{r^2}$ where $g(y) = y - (1+y)\log(1+y)$ for $y > 0$. Since $\frac{dg(y)}{dy} = -\log(1+y) < 0$, $f(y) \le f(0) = 0$. We thus have $\frac{d\exp_r(x)}{dx} \le 0$. □

We obtain the **reverse Young inequality with *r*-exponential function** which is a one-parameter extension of Theorem 2.5.1.

Corollary 2.5.1. *For* $t > 0$, $v \in [0,1]$ *and* $r \in (0,1]$,

$$\frac{(1-v)+vt}{t^v} \le \exp_r\left(v(1-v)\frac{(t-1)^2}{t}\right). \tag{2.5.13}$$

Proof. Since $\exp_1(v(1-v)\frac{(t-1)^2}{t}) = 1 + v(1-v)\frac{(t-1)^2}{t}$, we have the desired result by Theorem 2.5.3 and Lemma 2.5.4. □

Remark 2.5.4. Since $\lim_{r\to 0}\exp_r(x) = \exp(x)$ and Lemma 2.5.4, we have

$$1 + v(1-v)\frac{(t-1)^2}{t} \le \exp\left(v(1-v)\frac{(t-1)^2}{t}\right)$$

which means that the right-hand side in Theorem 2.5.3 gives the tighter upper bounds of $\frac{(1-v)+vt}{t^v}$ than one in Theorem 2.5.1.

Remark 2.5.5. Proposition 2.5.1 shows the upper bound of $\frac{(1-v)+vt}{t^v}$ is $M_v(t)$ for $0 < t \le 1$, while Theorem 2.5.3 gives the upper bound of $\frac{(1-v)+vt}{t^v}$ for all $t > 0$. In addition, for the case $t^v \le \frac{1}{2}$, the right-hand side in Theorem 2.5.3 gives the tighter upper bounds of $\frac{(1-v)+vt}{t^v}$ than $M_v(t)$ in Proposition 2.5.1.

Remark 2.5.6. It is easy to see that $1 + v(1-v)\frac{(t-1)^2}{t} \le 1 + \frac{(t-1)^2}{4t} = K(t)$. As we noted in Remark 2.5.1, the inequalities (2.5.5) are known. By the numerical computations, we have no ordering between $K^R(t)$ and $1 + v(1-v)\frac{(t-1)^2}{t}$. See [71] for the details.

Remark 2.5.7. From the proof of Lemma 2.5.4, we find that the function $\exp_r(x)$ is monotone decreasing for $r > 0$. However, the following inequality does not hold in general:

$$\frac{(1-v)+vt}{t^v} \le \exp_r\left(v(1-v)\frac{(t-1)^2}{t}\right), \quad (r > 1)$$

since we have a counterexample. See [71] for the details.

As similar way to the above, we can improve Theorem 2.5.2 in the following.

Theorem 2.5.4 ([71, Theorem 3.8]). *Let $t > 0$ and $v \in [0,1]$.*
(i) *If $0 < t \le 1$, then*

$$\frac{1}{1 - \frac{v(1-v)}{2}(t-1)^2} \le \frac{(1-v)+vt}{t^v} \le 1 + \frac{v(1-v)}{2}\left(\frac{1}{t}-1\right)^2.$$

(ii) *If $t \ge 1$, then*

$$\frac{1}{1 - \frac{v(1-v)}{2}(\frac{1}{t}-1)^2} \le \frac{(1-v)+vt}{t^v} \le 1 + \frac{v(1-v)}{2}(t-1)^2.$$

Proof. First, we prove the second inequality in (i). To this end, we set the function as

$$f_v(t) = 1 + \frac{v(1-v)}{2}\left(\frac{1}{t}-1\right)^2 - \frac{(1-v)+vt}{t^v}$$

for $0 < t \le 1$ and $0 \le v \le 1$. Then we find by elementary calculation

$$\frac{df_v(t)}{dt} = v(1-v)(t-1)t^{v+3}\left(t^v - t^2\right) \le 0$$

so that $f_v(t) \ge f_v(1) = 0$. Second, we prove the first inequality in (i). To this end, we set the function as

$$g_v(t) = \left\{(1-v) + vt\right\}\left\{1 - \frac{v(1-v)}{2}(t-1)^2\right\} - t^v$$

for $0 < t \le 1$ and $0 \le v \le 1$. Then we find by elementary calculations

$$\frac{dg_v(t)}{dt} = v\left\{-2t^{v-1} - 3v(1-v)t^2 - 2(3v^2 - 4v + 1)t + (3v^2 - 5v + 4)\right\}$$
$$\frac{d^2g_v(t)}{dt^2} = v(1-v)\left\{t^{v-2} - 3vt + (3v-1)\right\}, \quad \frac{d^3g_v(t)}{dt^3} = v(1-v)\left\{(v-2)t^{v-3} - 3v\right\} \le 0.$$

Thus we have $\frac{d^2g_v(t)}{dt^2} \ge \frac{d^2g_v(1)}{dt^2} = 0$ so that $\frac{dg_v(t)}{dt} \le \frac{dg_v(1)}{dt} = 0$ which implies $f_v(t) \ge f_v(1) = 0$.

By putting $t = 1/s$ in (i) and replacing v with $1 - v$, we obtain (ii). □

Lemma 2.5.5. *The function* $\exp_r(x)$ *defined for* $0 \le x \le 1$ *and* $-1 \le r < 0$*, is monotone decreasing in* r.

Proof. We calculate $\frac{d\exp_r(x)}{dx} = \frac{(1+rx)^{\frac{1-r}{r}} g(rx)}{r^2}$ where $g(y) = y - (1+y)\log(1+y)$ for $-1 \le y \le 0$. Since $\frac{dg(y)}{dy} = -\log(1+y) \ge 0$, $f(y) \le f(0) = 0$. We thus have $\frac{d\exp_r(x)}{dx} \le 0$. □

We obtain the **refined and reverse inequalities with r-exponential function** which are one-parameter extensions of Theorem 2.5.1.

Corollary 2.5.2 ([71, Corollary 3.10]). *Let* $t > 0$, $v \in [0,1]$, $r_1 \in [-1,0)$ *and* $r_2 \in (0,1]$.

(i) *If* $0 < t \le 1$*, then*

$$\exp_{r_1}\left(\frac{v(1-v)}{2}(t-1)^2\right) \le \frac{(1-v)+vt}{t^v} \le \exp_{r_2}\left(\frac{v(1-v)}{2}\left(\frac{1}{t}-1\right)^2\right).$$

(ii) *If* $t \ge 1$*, then*

$$\exp_{r_1}\left(\frac{v(1-v)}{2}\left(\frac{1}{t}-1\right)^2\right) \le \frac{(1-v)+vt}{t^v} \le \exp_{r_2}\left(\frac{v(1-v)}{2}(t-1)^2\right).$$

Proof. Taking into account $\exp_1(x) = 1 + x$, and applying two second inequalities in (i) and (ii) of Theorem 2.5.4 and Lemma 2.5.4, we obtain two second inequalities in (i) and (ii) of this theorem.

Taking into account $\exp_{-1}(x) = \frac{1}{1-x}$ and $\frac{v(1-v)}{2}(t-1)^2 < 1$ for $0 < t \le 1$, $\frac{v(1-v)}{2}(\frac{1}{t} - 1)^2 < 1$ for $t \ge 1$, and applying two first inequalities in (i) and (ii) of Theorem 2.5.4 and Lemma 2.5.5, we obtain two first inequalities in (i) and (ii) of this corollary. □

Remark 2.5.8. As we noted $\lim_{r\to 0} \exp_r(x) = \exp(x)$, and Lemma 2.5.4 and Lemma 2.5.5 assure that Theorem 2.5.4 gives tighter bounds of $\frac{(1-v)+vt}{t^v}$ than Theorem 2.5.2.

Remark 2.5.9. Bounds in (i) of Theorem 2.5.4 can be compared with Proposition 2.5.1. As for upper bound, $M_v(t)$ is tighter than the right-hand side in the second inequality of (i) of Theorem 2.5.4. Since $(\frac{t+1}{2})^{v+1} \geq t^2$ for $0 < t \leq 1$, $m_v(t)$ is also tighter than the left-hand side in the first inequality of (i) of Theorem 2.5.4. The inequality $(\frac{t+1}{2})^{v+1} \geq t^2$ can be proven in the following. We set $f_v(t) = (v+1)\log\frac{t+1}{2} - 2\log t$, then we have $\frac{df_v(t)}{dt} = \frac{(v-1)t-2}{t(t+1)} \leq 0$ so that $f_v(t) \geq f_v(1) = 0$.

Note that Remark 2.5.9 and (i) of Corollary 2.5.2 give Proposition 2.5.2. However, Corollary 2.5.2 gives alternative tight bounds of $\frac{(1-v)+vt}{t^v}$ when $t \geq 1$.

Remark 2.5.10. We give comparisons our inequalities obtained in Theorem 2.5.4 with the inequalities (2.5.5). By the numerical computations, we have no ordering between $K^R(t)$ and $1 + v(1-v)\frac{(t-1)^2}{2t^2}$ for $0 < t \leq 1$. See [71] for the details.

We close this subsection showing operator versions for Corollary 2.5.1 and Corollary 2.5.2 in the following.

Corollary 2.5.3. *Let* $0 \leq v \leq 1$, $0 < r \leq 1$ *and let* A *and* B *be strictly positive operators satisfying (i)* $0 < m \leq A \leq m' < M' \leq B \leq M$ *or (ii)* $0 < m \leq B \leq m' < M' \leq A \leq M$ *with* $h = \frac{M}{m}$. *Then*

$$A\nabla_v B \leq \exp_r\big(4v(1-v)(K(h)-1)\big)A\sharp_v B.$$

Corollary 2.5.4. *Let* $0 \leq v \leq 1$, $-1 \leq r_1 < 0$, $0 < r_2 \leq 1$ *and let* A *and* B *be strictly positive operators satisfying (i)* $0 < m \leq A \leq m' < M' \leq B \leq M$ *or (ii)* $0 < m \leq B \leq m' < M' \leq A \leq M$ *with* $h = \frac{M}{m}$ *and* $h' = \frac{M'}{m'}$. *Then*

$$\exp_{r_1}\left(\frac{v(1-v)}{2}\left(\frac{h-1}{h}\right)^2\right)A\sharp_v B \leq A\nabla_v B \leq \exp_{r_2}\left(\frac{v(1-v)}{2}(h'-1)^2\right)A\sharp_v B.$$

Remark 2.5.11. In [243, Theorem 6 (i)], L. Zhu obtained the inequality

$$\frac{(1-v)+vt}{t^v} \leq 1 + \frac{v(1-v)}{2}\frac{(t-1)^2}{t}, \quad (1/2 \leq v \leq 1, t > 0)$$

which refines (2.5.12) and the second inequalities in Theorem 2.5.4 for the case of $1/2 \leq v \leq 1$. In addition, the refinements of the first inequalities in Theorem 2.5.4 were given in [243, Theorem 8]:

$$\frac{1}{1 - \frac{v(1-v)}{2}(2(\frac{1}{\sqrt{t}}-1))^2} \leq \frac{(1-v)+vt}{t^v}, \quad (0 \leq v \leq 1, t \geq 1).$$

The case $0 < t \leq 1$ can be also obtained by putting $t = 1/s$ and replacing v with $1-v$.

2.6 Young inequalities with new constants

Based on the refined scalar Young inequality, F. Kittaneh and Y. Manasrah [131] obtained the operator inequalities

$$2r(A\nabla B - A\sharp B) \le A\nabla_v B - A\sharp_v B \le 2R(A\nabla B - A\sharp B), \tag{2.6.1}$$

where $r = \min\{v, 1-v\}$ and $R = \max\{v, 1-v\}$. H. Zuo et al. [248] refined operator Young inequality with the Kantorovich constant $K(h)$, $(h > 0)$, and showed the following inequality:

$$A\nabla_v B \ge K^r(h) A\sharp_v B, \tag{2.6.2}$$

where $0 < \alpha' I \le A \le \alpha I \le \beta I \le B \le \beta' I$ or $0 < \alpha' I \le B \le \alpha I \le \beta I \le A \le \beta' I$ with $h = \frac{\beta}{\alpha}$. Note also that the inequality (2.6.2) improves Furuichi's result given in the first inequality of Theorem 2.2.2 shown in [65].

As for the reverse of the operator Young inequality, under the same conditions, W. Liao et al. [140] gave the following inequality:

$$A\nabla_v B \le K^R(h') A\sharp_v B, \tag{2.6.3}$$

where $0 < \alpha' I \le A \le \alpha I \le \beta I \le B \le \beta' I$ or $0 < \alpha' I \le B \le \alpha I \le \beta I \le A \le \beta' I$ with $h' = \frac{\beta'}{\alpha'}$ Two inequalities (2.6.2) and (2.6.3) with Kantorovich constant:

$$K^r(h) A\sharp_v B \le A\nabla_v B \le K^R(h') A\sharp_v B$$

give a refinement and reverse for operator Young inequality. We call the above inequalities **Zuo–Liao inequalities** for short.

This section intends to give some **operator Young inequalities with new constants** via the Hermite–Hadamard inequality. Our new constants are different from Kantorovich constant shown in the Zuo–Liao inequality. That is, the following theorem is one of the main results in this section.

Theorem 2.6.1 ([79]). *Let A, B be strictly positive operators such that $0 < h' I \le A^{-\frac{1}{2}} B A^{-\frac{1}{2}} \le hI \le I$ for some positive scalars h and h'. Then for each $0 \le v \le 1$,*

$$m_v(h) A\sharp_v B \le A\nabla_v B \le M_v(h') A\sharp_v B, \tag{2.6.4}$$

where

$$m_v(x) = 1 + \frac{2^v v(1-v)(x-1)^2}{(x+1)^{v+1}}, \quad M_v(x) = 1 + \frac{v(1-v)(x-1)^2}{2x^{v+1}}.$$

The inequalities in (2.6.4) give alternative Zuo–Liao inequalities with new constants. In order to prove our results, we need the well-known **Hermite–Hadamard**

inequality which asserts that if $f : [a,b] \to \mathbb{R}$ is a convex function, then the following chain of inequalities hold:

$$f\left(\frac{a+b}{2}\right) \le \frac{1}{b-a}\int_a^b f(x)\,dx \le \frac{f(a)+f(b)}{2}. \tag{2.6.5}$$

If f is a concave function, then the signs in inequalities (2.6.5) are reversed.

Our first attempt, which is a direct consequence of [176, Theorem 1], gives a difference refinement and reverse for the operator Young inequality via (2.6.5).

The next lemma provides a technical result which we will need in the sequel.

Lemma 2.6.1. *Let $v \in (0,1]$.*
(i) *For each $t > 0$, the function $f_v(t) = v(1-t^{v-1})$ is concave.*
(ii) *The function $g_v(t) = \frac{v(1-v)(t-1)}{t^{v+1}}$, is concave if $t \le 1 + \frac{2}{v}$, and convex if $t \ge 1 + \frac{2}{v}$.*

Proof. The function $f_v(t)$ is twice differentiable and $f_v''(t) = v(1-v)(v-2)t^{v-3}$. According to the assumptions $t > 0$ and $0 \le v \le 1$, we have $f_v''(t) \le 0$.

The function $g_v(t)$ is also twice differentiable and $g_v''(t) = v(1-v)(v+1)(\frac{vt-v-2}{t^{v+3}})$ which implies (ii). □

Using this lemma, together with (2.6.5), we have the following proposition.

Proposition 2.6.1. *Let A, B be strictly positive operators such that $A \le B$. Then for each $0 \le v \le 1$,*

$$\begin{aligned} v(B-A)A^{-1}\left(\frac{A - A\natural_{v-1}B}{2}\right) &\le A\nabla_v B - A\sharp_v B \\ &\le v(B-A)A^{-1}\left(A - A^{\frac{1}{2}}\left(\frac{I + A^{-\frac{1}{2}}BA^{-\frac{1}{2}}}{2}\right)^{v-1}A^{\frac{1}{2}}\right). \end{aligned} \tag{2.6.6}$$

By virtue of Proposition 2.6.1, we can improve the first inequality in (2.0.3).

Remark 2.6.1. It is worth remarking that the left-hand side of inequality (2.6.6), is a refinement of operator Young inequality in the sense of $v(x-1)(\frac{1-x^{v-1}}{2}) \ge 0$ for each $x \ge 1$ and $0 \le v \le 1$. Replacing A and B by A^{-1} and B^{-1} respectively and taking inverse, we get

$$\begin{aligned} A!_v B &\le \left\{v(B^{-1} - A^{-1})A\left(\frac{A^{-1} - A^{-1}\natural_{v-1}B^{-1}}{2}\right) + A^{-1}\sharp_v B^{-1}\right\}^{-1} \\ &\le A\sharp_v B \\ &\le v(B-A)A^{-1}\left(\frac{A - A\natural_{v-1}B}{2}\right) + A\sharp_v B \\ &\le A\nabla_v B. \end{aligned} \tag{2.6.7}$$

We easily find that the inequalities in (2.6.7) provide the interpolation among arithmetic mean, geometric mean and harmonic mean. That is, (2.6.7) gives a difference refinement for operator Young inequalities.

In order to give a proof of our first main result in this section, we need the following essential result.

Proposition 2.6.2. *For each* $0 < x \leq 1$ *and* $0 \leq v \leq 1$, *the functions* $m_v(x)$ *and* $M_v(x)$ *defined in Theorem* 2.6.1 *are decreasing. Moreover,* $1 \leq m_v(x) \leq M_v(x)$.

We are now in a position to prove Theorem 2.6.1 which are **operator Young inequalities with new constants**.

Proof of Theorem 2.6.1. It is routine to check that the function $f_v(t) = \frac{v(1-v)(t-1)}{t^{v+1}}$ where $0 < t \leq 1$ and $v \in [0,1]$, is concave. We can verify that

$$\int_x^1 f_v(t)\,dt = 1 - \frac{(1-v) + vx}{x^v}.$$

Hence from the inequality (2.6.5) we can write

$$m_v(x)x^v \leq (1-v) + vx \leq M_v(x)x^v, \tag{2.6.8}$$

for each $0 < x \leq 1$ and $v \in [0,1]$.

Now, we shall use the same procedure as in [65, Theorem 2]. The inequality (2.6.8) implies that

$$\min_{h' \leq x \leq h \leq 1} m_v(x)x^v \leq (1-v) + vx \leq \max_{h' \leq x \leq h \leq 1} M_v(x)x^v.$$

Based on this inequality, one can see easily for which X,

$$\min_{h' \leq x \leq h \leq 1} m_v(x)X^v \leq (1-v)I + vX \leq \max_{h' \leq x \leq h \leq 1} M_v(x)X^v. \tag{2.6.9}$$

By substituting $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$ for X and taking into account that $m_v(x)$ and $M_v(x)$ are decreasing, the relation (2.6.9) implies

$$m_v(h)\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^v \leq (1-v)I + vA^{-\frac{1}{2}}BA^{-\frac{1}{2}} \leq M_v(h')\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^v. \tag{2.6.10}$$

Multiplying $A^{\frac{1}{2}}$ from the both sides to the inequality (2.6.10), we have the inequality (2.6.4). □

Remark 2.6.2. Notice that, the condition $0 < h'I \leq A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \leq hI \leq I$ in Theorem 2.6.1, can be replaced by $0 < \alpha' I \leq B \leq \alpha I \leq \beta I \leq A \leq \beta' I$. In this case, we have

$$m_v(h)A\sharp_v B \leq A\nabla_v B \leq M_v(h')A\sharp_v B,$$

where $h = \frac{\alpha}{\beta}$ and $h' = \frac{\alpha'}{\beta'}$.

Theorem 2.6.1 can be used to infer the following remark.

Remark 2.6.3. Assume the conditions of Theorem 2.6.1. Then

$$m_v(h)A!_vB \le A\sharp_vB \le M_v(h')A!_vB.$$

The left-hand side of the inequality (2.6.4) can be squared. In order to prove the following corollary, we prepare two lemmas which will be also applied in the later chapters.

Lemma 2.6.2 ([25, Theorem 1]). *For $A, B > 0$, we have $\|AB\| \le \frac{1}{4}\|A+B\|^2$.*

Lemma 2.6.3 (Choi inequality [22, p. 41]). *Let $A > 0$. Then for every normalized positive linear map Φ,*

$$\Phi^{-1}(A) \le \Phi(A^{-1}).$$

Corollary 2.6.1. *Let $0 < \alpha' I \le B \le \alpha I \le \beta I \le A \le \beta' I$. Then for every normalized positive linear map Φ,*

$$\Phi^2(A\nabla_vB) \le \left(\frac{K(h')}{m_v(h)}\right)^2 \Phi^2(A\sharp_vB) \tag{2.6.11}$$

and

$$\Phi^2(A\nabla_vB) \le \left(\frac{K(h')}{m_v(h)}\right)^2 (\Phi(A)\sharp_v\Phi(B))^2, \tag{2.6.12}$$

where $h = \frac{\alpha}{\beta}$ and $h' = \frac{\alpha'}{\beta'}$.

Proof. According to the assumptions,

$$(\alpha' + \beta')I \ge \alpha'\beta' A^{-1} + A, \qquad (\alpha' + \beta')I \ge \alpha'\beta' B^{-1} + B,$$

since $(t-\alpha')(t-\beta') \le 0$ for $\alpha' \le t \le \beta'$. From these, we can write

$$(\alpha' + \beta')I \ge \alpha'\beta'\Phi(A^{-1}\nabla_vB^{-1}) + \Phi(A\nabla_vB), \tag{2.6.13}$$

where Φ is a normalized positive linear map. We have

$$\begin{aligned}
&\|\Phi(A\nabla_vB)\alpha'\beta' m_v(h)\Phi^{-1}(A\sharp_vB)\| \\
&\le \frac{1}{4}\|\Phi(A\nabla_vB) + \alpha'\beta' m_v(h)\Phi^{-1}(A\sharp_vB)\|^2 \quad \text{(by Lemma 2.6.2)} \\
&\le \frac{1}{4}\|\Phi(A\nabla_vB) + \alpha'\beta' m_v(h)\Phi(A^{-1}\sharp_vB^{-1})\|^2 \quad \text{(by Lemma 2.6.3)} \\
&\le \frac{1}{4}\|\Phi(A\nabla_vB) + \alpha'\beta'\Phi(A^{-1}\nabla_vB^{-1})\|^2 \quad \text{(by Remark 2.6.2)} \\
&\le \frac{1}{4}(\alpha' + \beta')^2 \quad \text{(by (2.6.13))}.
\end{aligned}$$

This is equivalent to

$$\|\Phi(A\nabla_v B)\Phi^{-1}(A\sharp_v B)\| \le \frac{K(h')}{m_v(h)}, \tag{2.6.14}$$

where $h = \frac{\alpha}{\beta}$ and $h' = \frac{\alpha'}{\beta'}$. It is not hard to see that (2.6.14) is equivalent to (2.6.11). The proof of the inequality (2.6.12) goes likewise and we omit the details. □

Remark 2.6.4. Obviously, the bounds in (2.6.11) and (2.6.12) are tighter than those in [142, Theorem 2.1], under the conditions $0 < \alpha' I \le B \le \alpha I \le \beta I \le A \le \beta' I$ with $h = \frac{\alpha}{\beta}$ and $h' = \frac{\alpha'}{\beta'}$.

We here point out connections between our results given in the above and some inequalities proved in other contexts. That is, we are now going to explain the advantages of our results. Let $v \in [0,1]$, $r = \min\{v, 1-v\}$, $R = \max\{v, 1-v\}$ and $m_v(\cdot)$, $M_v(\cdot)$ were defined as in Theorem 2.6.1. As we will show in the below, the following proposition explains the advantages of our results (Proposition 2.6.1 and Theorem 2.6.1).

Proposition 2.6.3. *The following statements are true:*

(I-i) *The lower bound of Proposition* 2.6.1 *improves the first inequality in* (2.6.1)*, when* $\frac{3}{4} \le v \le 1$ *with* $0 < A \le B$.

(I-ii) *The upper bound of Proposition* 2.6.1 *improves the second inequality in* (2.6.1)*, when* $\frac{2}{3} \le v \le 1$ *with* $0 < A \le B$.

(I-iii) *The upper bound of Proposition* 2.6.1 *improves the second inequality in* (2.6.1)*, when* $0 \le v \le \frac{1}{3}$ *with* $0 < A \le B$.

(II) *The upper bound of Theorem* 2.6.1 *improves the inequality*

$$(1-v) + vx \le x^v K(x),$$

when $x^v \ge \frac{1}{2}$.

(III) *The upper bound of Theorem* 2.6.1 *improves the inequality given by S. S. Dragomir in* [43*, Theorem 1*]*,*

$$(1-v) + vx \le \exp\big(4v(1-v)(K(x)-1)\big)x^v, \quad x > 0 \tag{2.6.15}$$

when $0 \le v \le \frac{1}{2}$ *and* $0 < x \le 1$.

(IV) *There is no ordering between Theorem* 2.6.1 *and the inequalities (*2.6.2*) and (*2.6.3*).*

Therefore, we conclude that Proposition 2.6.1 and Theorem 2.6.1 are not trivial results. Comparisons with known results, and proofs are given in [79].

2.7 More accurate inequalities

In this section, we prove the *operator mean inequality*

$$A\nabla_v B \le \frac{m\nabla_r M}{m\sharp_r M} A\sharp_v B, \quad r = \min\{v, 1-v\}, \tag{2.7.1}$$

when $0 < mI \le A, B \le MI$ and $0 \le v \le 1$. This is an improvement of the inequality by M. Tominaga [226],

$$A\nabla_v B \le S(h) A\sharp_v B$$

(it should be noticed here that the original proof of this inequality is shown in [4, Theorem 1]), and the inequality given by W. Liao et al. [140],

$$A\nabla_v B \le K^R(h) A\sharp_v B, \quad R = \max\{v, 1-v\}, \tag{2.7.2}$$

where $h = \frac{M}{m}$ for $0 < mI \le A, B \le MI$. We here give the proof of the inequality (2.7.1).

Proof of the inequality (2.7.1). We first prove the corresponding scalar inequality. Define

$$f(x) = \frac{(1-v) + vx}{x^v}, \quad v \in [0, 1].$$

We aim to find the maximum of f on $[\frac{m}{M}, \frac{M}{m}]$. Direct calculations show that

$$f'(x) = \frac{x^{v-1}(x-1)(v-v^2)}{x^{2v}}.$$

Therefore, f attains its absolute maximum at $x = \frac{m}{M}$ or $x = \frac{M}{m}$. To compare between $f(\frac{m}{M})$ and $f(\frac{M}{m})$, let

$$g(h) = \frac{(1-v) + vh}{h^v} - \frac{(1-v)h + v}{h^{1-v}}, \quad h > 1 \text{ and } v \in [0, 1].$$

Then

$$g'(h) = (v-1)v\left(\frac{h-1}{h^{v+2}}\right)(h^{2v} - h),$$

and $g(1) = 0$. It is clear that for $v \in [0, \frac{1}{2}]$ we have $g'(h) \ge 0$, while $g'(h) \le 0$ when $v \in [\frac{1}{2}, 1]$. Therefore, when $v \in [0, \frac{1}{2}]$, we have $g(h) \ge g(1) = 0$, while $g(h) \le 0$ when $v \in [\frac{1}{2}, 1]$. Since $\frac{M}{m} > 1$, it follows that

$$\max_{x \in [\frac{m}{M}, \frac{M}{m}]} f(x) = \begin{cases} \frac{m\nabla_v M}{m\sharp_v M} & \text{for } 0 \le v \le \frac{1}{2} \\ \frac{M\nabla_v m}{M\sharp_v m} & \text{for } \frac{1}{2} \le v \le 1. \end{cases} \tag{2.7.3}$$

Then (2.7.3) says that

$$\begin{cases} (1-v) + vx \le \frac{m\nabla_v M}{m\sharp_v M} x^v & \text{for } 0 \le v \le \frac{1}{2} \\ (1-v) + vx \le \frac{M\nabla_v m}{M\sharp_v m} x^v & \text{for } \frac{1}{2} \le v \le 1, \end{cases}$$

which is equivalent to saying

$$(1-v)+vx \le \frac{m\nabla_r M}{m\sharp_r M} x^v,$$

where $r = \min\{v, 1-v\}$, $v \in [0,1]$ and $\frac{m}{M} \le x \le \frac{M}{m}$. Now, if $mI \le A, B \le MI$, it follows that $\frac{m}{M} \le A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le \frac{M}{m}$ and a standard functional calculus applied to the numerical inequality implies (2.7.1). □

The constant $\frac{m\nabla_r M}{m\sharp_r M}$ is best possible in (2.7.1) in the sense that smaller quantities cannot replace it. As a matter of fact, let $A = mI$ and $B = MI$. Then we have

$$A\nabla_v B = (m\nabla_v M)I \quad \text{and} \quad A\sharp_v B = (m\sharp_v M)I.$$

Then it is immediate to see that

$$A\nabla_v B = \frac{m\nabla_v M}{m\sharp_v M} A\sharp_v B, \quad v \in \left[0, \frac{1}{2}\right]$$

as desired. In addition, by choosing $A = MI$ and $B = mI$, we get

$$A\nabla_v B = (M\nabla_v m)I \quad \text{and} \quad A\sharp_v B = (M\sharp_v m)I.$$

Therefore,

$$A\nabla_v B = \frac{M\nabla_v m}{M\sharp_v m} A\sharp_v B, \quad v \in \left[\frac{1}{2}, 1\right].$$

In the next proposition, we show that our estimate in (2.7.1) is better than the inequality (2.7.2) given by W. Liao et al. [140].

Proposition 2.7.1. *Let $m, M > 0$, $v \in [0,1]$ and $r = \min\{v, 1-v\}$. Then*

$$\frac{m\nabla_r M}{m\sharp_r M} \le \left(\frac{m\nabla M}{m\sharp M}\right)^{2(1-r)}. \tag{2.7.4}$$

Proof. We see that the inequality (2.7.2) is equivalent to saying

$$A\nabla_v B \le \left(\frac{m\nabla M}{m\sharp M}\right)^{2(1-r)} A\sharp_v B. \tag{2.7.5}$$

First, we assume that $v \in [0, \frac{1}{2}]$. Let

$$g(v) = \log(1-v+vM) - v\log M - 2(1-v)\{\log(1+M) - \log(2\sqrt{M})\}.$$

Then direct computations show that

$$g''(v) = -\frac{(M-1)^2}{(1+(M-1)v)^2} \le 0.$$

Hence, g' is decreasing and

$$g'(v) \geq g'\left(\frac{1}{2}\right) = 2\left(\log(M+1) - \log M + \frac{M-1}{M+1} - \log 2\right).$$

Letting $h(M) = \log(M+1) - \log M + \frac{M-1}{M+1} - \log 2$, we have $h'(M) = \frac{2(M-1)}{M(M+1)^2}$. Therefore, h attains its minimum at $M = 1$ and $h(M) \geq h(1) = 0$. Consequently, $g' \geq 0$ and g is increasing on $[0, \frac{1}{2}]$. Hence, for $v \in [0, \frac{1}{2}]$, we have $g(v) \leq g(\frac{1}{2}) = 0$. Replacing M by $\frac{M}{m}$ and rearranging g, we have

$$\frac{m\nabla_v M}{m\sharp_v M} \leq \left(\frac{m\nabla M}{m\sharp M}\right)^{2(1-v)}, \quad v \in [0, 1/2].$$

For $v \in [\frac{1}{2}, 1]$, let $k(v) = g(1-v)$. Since $g = g(v)$ is increasing on $[0, \frac{1}{2}]$, k is decreasing on $[\frac{1}{2}, 1]$. So, $k(v) \leq k(\frac{1}{2})$ which implies the required inequality for $v \in [\frac{1}{2}, 1]$. □

By virtue of the inequality (2.7.1), we have the following comparison between the geometric and harmonic means:

$$A\sharp_v B \leq \frac{m\nabla_r M}{m\sharp_r M} A!_v B. \tag{2.7.6}$$

Further, we have the following double-sided inequality, for $v \in [0, 1]$:

$$A\nabla_v B - M\left(\frac{m\nabla_r M}{m\sharp_r M} - 1\right)I \leq A\sharp_v B \leq M\left(\frac{m\nabla_r M}{m\sharp_r M} - 1\right)I + A!_v B, \tag{2.7.7}$$

where $r = \min\{v, 1-v\}$. The inequalities (2.7.7) are difference reverses for operator Young inequalities given in (2.0.3).

The inequality (2.7.1) can be also used to present the following Hilbert–Schmidt norm (Frobenius norm) versions, for the algebra $M(n, \mathbb{C})$ of all complex matrices of size $n \times n$. Recall that for positive matrices A, B and an arbitrary $X \in M(n, \mathbb{C})$, we have the Young-type inequality [133]

$$\|A^{1-v} X B^v\|_2 \leq \|(1-v)AX + vXB\|_2, \quad v \in [0, 1]. \tag{2.7.8}$$

It is interesting that the inequality (2.7.8) is not valid for arbitrary unitarily invariant norms. Now we present a reversed version of (2.7.8).

Corollary 2.7.1. *Let $A, B \in M_+(n, \mathbb{C})$ and $X \in M(n, \mathbb{C})$. If $mI \leq A, B \leq MI$ with $0 < m < M$, then we have*

$$\|(1-v)AX + vXB\|_2 \leq \frac{m\nabla_r M}{m\sharp_r M} \|A^{1-v} X B^v\|_2, \tag{2.7.9}$$

where $v \in [0, 1]$ and $r = \min\{v, 1-v\}$.

Proof. Let U and V be unitary matrices such that

$$A = U \operatorname{diag}(\lambda_i) U^* \quad \text{and} \quad B = V \operatorname{diag}(\mu_i) V^*$$

are spectral decompositions of A and B. Moreover, let $Y = U^*XV$. Then

$$\begin{aligned}\|(1-v)AX + vXB\|_2^2 &= \|U((1-v)\operatorname{diag}(\lambda_i)Y + vY\operatorname{diag}(\mu_i))V^*\|_2^2 \\ &= \|[(1-v)\lambda_i + v\mu_j] \circ [y_{ij}]\|_2^2 \\ &= \sum_{i,j=1}^{n} ((1-v)\lambda_i + v\mu_j)^2 |y_{ij}|^2 \\ &\le \left(\frac{m\nabla_r M}{m\sharp_r M}\right)^2 \sum_{i,j=1}^{n} (\lambda_i^{1-v}\mu_j^{v})^2 |y_{ij}|^2 \\ &= \left(\frac{m\nabla_r M}{m\sharp_r M}\right)^2 \|A^{1-v}XB^{v}\|_2^2,\end{aligned}$$

which completes the proof. □

On the other hand, a **reverse Heinz inequality** can be also found as follows. Recall that for positive A, B and arbitrary $X \in M(n,\mathbb{C})$ we have [23]

$$\|A^{1-v}XB^{v} + A^{v}XB^{1-v}\|_u \le \|AX + XB\|_u, \quad v \in [0,1], \tag{2.7.10}$$

for any unitarily invariant norm $\|\cdot\|_u$. Now we present a reversed version of this inequality for the Hilbert–Schmidt norm.

Corollary 2.7.2. *Let $A, B \in M_+(n,\mathbb{C})$ and $X \in M(n,\mathbb{C})$. If $mI \le A,\ B \le MI$ with $0 < m < M$, then we have*

$$\|AX + XB\|_2 \le \frac{m\nabla_r M}{m\sharp_r M}\|A^{1-v}XB^{v} + A^{v}XB^{1-v}\|_2, \tag{2.7.11}$$

where $v \in [0,1]$ and $r = \min\{v, 1-v\}$.

Proof. Notice first that if a and b are positive numbers in $[m, M]$, then

$$a\nabla_v b \le \frac{m\nabla_r M}{m\sharp_r M} a\sharp_v b \quad \text{and} \quad a\nabla_{1-v} b \le \frac{m\nabla_r M}{m\sharp_r M} a\sharp_{1-v} b,$$

where $r = \min\{v, 1-v\}$. Adding these inequalities implies

$$a + b \le \frac{m\nabla_r M}{m\sharp_r M}(a^{1-v}b^{v} + a^{v}b^{1-v}).$$

Then a standard argument like that of Corollary 2.7.1 implies the required inequality. □

At this stage, it is quite natural to ask whether inequalities (2.7.9) and (2.7.11) are valid for an arbitrary unitarily invariant norm. Numerical examples support this claim, but we do not have a definite answer.

Remark 2.7.1. Notice that when $mI \le A, B \le MI$ with $m < M$, then $m'I \le A, B \le M'I$, when $m' \le m$ and $M' \ge M$. Therefore, we have

$$A\nabla_v B \le \frac{m'\nabla_r M'}{m'\sharp_r M'} A\sharp_v B, \quad r = \min\{v, 1-v\},$$

giving an infinite set of inequalities. However, it turns out that the best inequality is attained when m and M are the largest and smallest such numbers, respectively. This is due to the following simple observation.

Proposition 2.7.2. *Let v be a fixed number in $[0,1]$. If $M > 0$ is fixed, then the mapping $m \mapsto \frac{m\nabla_r M}{m\sharp_r M}$ is decreasing on $(0, M]$. On the other hand, if $m > 0$ is fixed, then the mapping $M \mapsto \frac{m\nabla_r M}{m\sharp_r M}$ is increasing on $[m, \infty)$.*

Therefore, the minimum value of the constant $\frac{m\nabla_r M}{m\sharp_r M}$ is attained when $m = \max\{a > 0 : aI \le A, B\}$ and $M = \min\{b > 0 : A, B \le bI\}$.

2.7.1 Inequalities for positive linear maps

In this subsection, we present inequalities that govern positive linear maps and their squared versions. The problem of squaring operator inequalities has been studied extensively in the literature. To our surprise, M. Lin [142, Theorem 2.1] showed that the reverse Young inequality can be squared. His method was based on some observations about the operator norm and Young inequality of R. Bhatia and F. Kittaneh.

Our next target is to show squared versions of our inequalities involving the new constant $\frac{m\nabla_r M}{m\sharp_r M}$. For this purpose, we remind the reader of the following simple method that allows squaring such inequalities.

Lemma 2.7.1 ([57, Theorem 6]). *Let $A, B \in \mathbb{B}(\mathcal{H})$ such that $0 < A \le B$ and $0 < mI \le A \le MI$. Then*

$$A^2 \le K(h)B^2,$$

where $h = \frac{M}{m}$.

A direct application of Lemma 2.7.1 implies the following squared version of the reverse Young inequality filtered through a normalized positive linear map, providing an alternative proof of a main result in [142] by M. Lin.

Corollary 2.7.3. *Let $A, B \in \mathbb{B}(\mathcal{H})$ such that $0 < mI \le A, B \le MI$. Then for every normalized positive linear map Φ*

$$\Phi^2(A\nabla B) \le K^2(h)\Phi^2(A\sharp B), \tag{2.7.12}$$

and

$$\Phi^2(A\nabla B) \le K^2(h)\big(\Phi(A)\sharp\Phi(B)\big)^2, \tag{2.7.13}$$

where $h = \frac{M}{m}$.

Proof. From [53, Theorem 13], we have $A\nabla B \le \sqrt{K(h)}(A\sharp B)$. Since $mI \le A\nabla B \le MI$, by applying Lemma 2.7.1 we can get (2.7.12). The inequality (2.7.13) follows similarly from the inequality $\Phi(A\nabla B) = \Phi(A)\nabla\Phi(B) \le \sqrt{K(h)}(\Phi(A)\sharp\Phi(B))$. □

The significance of Lemma 2.7.1 is the way that it can be used to square some of the inequalities we have just proved. For example, we have the following squared version of (2.7.1), where the constant $(\frac{m\nabla_r M}{m\sharp_r M})^2$ appears.

Corollary 2.7.4. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be such that $mI \le A, B \le MI$ for some scalars $0 < m < M$. Then, for any normalized positive linear map Φ,*

$$\Phi^2(A\nabla_v B) \le K(h)\left(\frac{m\nabla_r M}{m\sharp_r M}\right)^2 \Phi^2(A\sharp_v B),$$

and

$$\Phi^2(A\nabla_v B) \le K(h)\left(\frac{m\nabla_r M}{m\sharp_r M}\right)^2 \left(\Phi(A)\sharp_v\Phi(B)\right)^2,$$

where $h = \frac{m}{M}$, $r = \min\{v, 1-v\}$ and $v \in [0, 1]$.

In particular, when $v = \frac{1}{2}$, we get Corollary 2.7.3.

Remark 2.7.2. If we apply similar argument as in Corollary 2.7.3, we can obtain the squared version of the reverse geometric-harmonic mean inequality utilizing (2.7.6).

Remark 2.7.3. As we have seen Corollaries 2.7.3 and 2.7.4, it is natural to consider the relation between $K(h)(\frac{m\nabla_r M}{m\sharp_r M})^2$ and $K^2(h)$ with $h = \frac{m}{M}$. However, by simple calculations there is no ordering between the quantities $\frac{m\nabla_r M}{m\sharp_r M}$ and $\frac{m\nabla M}{m\sharp M}$ in general.

Proposition 2.7.3. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be such that $mI \le A, B \le MI$ for some scalars $0 < m < M$. Then, for any normalized positive linear map Φ,*

$$\Phi^2(A\nabla_v B) \le C\Phi^2(A\sharp_v B), \quad \Phi^2(A\nabla_v B) \le C\left(\Phi(A)\sharp_v\Phi(B)\right)^2,$$

where $C = \min\{K(M/m)(\frac{m\nabla_r M}{m\sharp_r M})^2, K^2(M/m)\}$, $r = \min\{v, 1-v\}$ and $v \in [0, 1]$.

2.7.2 Inequalities to operator monotone functions

In the subsequent discussion, we consider operator means with operator monotone functions defined on the half real line $(0, \infty)$. The following lemma is well known. However, we give the proof to keep the present book self-contained.

Lemma 2.7.2. *Let $f : (0, \infty) \to (0, \infty)$ be a given function and let $\alpha \ge 1$.*

(i) *If f is operator monotone, then*

$$f(\alpha t) \le \alpha f(t). \tag{2.7.14}$$

(ii) *If g is operator monotone decreasing, then*

$$g(\alpha t) \geq \frac{1}{\alpha} g(t).$$

Proof.

(i) Since $f(t)$ is operator monotone, then $\frac{t}{f(t)}$ is also operator monotone [97, Corollary 1.14]. Hence for $\alpha \geq 1$,

$$\frac{t}{f(t)} \leq \frac{\alpha t}{f(\alpha t)} \Rightarrow \alpha f(t) \geq f(\alpha t).$$

(ii) If g is operator monotone decreasing, then $\frac{1}{g}$ is operator monotone. Applying inequality (2.7.14) for $f = \frac{1}{g}$, we infer

$$g(\alpha t)^{-1} \leq \alpha g(t)^{-1} \Rightarrow g(\alpha t) \geq \frac{1}{\alpha} g(t). \qquad \square$$

Now we are ready to present the following result comparing between $f(A)\sharp_v f(B)$ and $f(A\sharp_v B)$.

Theorem 2.7.1 ([104]). *Let $f : (0,\infty) \to (0,\infty)$ be an operator monotone function and $A, B \in \mathbb{B}(\mathcal{H})$ are such that $mI \leq A, B \leq MI$ for some scalars $0 < m < M$. Then*

$$f(A)\sharp_v f(B) \leq \frac{m\nabla_r M}{m\sharp_r M} f(A\sharp_v B), \tag{2.7.15}$$

where $r = \min\{v, 1-v\}$ and $v \in [0,1]$.

Proof. For the given parameters,

$$\begin{aligned} f(A)\sharp_v f(B) &\leq f(A)\nabla_v f(B) \quad \text{(by AM-GM inequality)} \\ &\leq f(A\nabla_v B) \quad \text{(by [97, Corollary 1.12])} \\ &\leq f\left(\frac{m\nabla_r M}{m\sharp_r M} A\sharp_v B\right) \quad \text{(by (2.7.1))} \\ &\leq \frac{m\nabla_r M}{m\sharp_r M} f(A\sharp_v B) \quad \text{(by Lemma 2.7.2 (i))}, \end{aligned} \tag{2.7.16}$$

which completes the proof. $\square$

Remark 2.7.4. Theorem 2.7.1 provides a refinement of [99, Corollary 1], because of the fact: $(1 \leq)\frac{m\nabla_r M}{m\sharp_r M} \leq S(\frac{M}{m})$. We prove it here for convenience. According to [226, Lemma 2.3], if $a, b > 0$ and $v \in [0,1]$, then

$$a\nabla_v b \leq S\left(\frac{a}{b}\right) a\sharp_v b, \tag{2.7.17}$$

Setting $a = M$ and $b = m$ in (2.7.17), we get $\frac{M\nabla_v m}{M\sharp_v m} \le S(\frac{M}{m})$. On the other hand, by substituting $a = m$ and $b = M$ in (2.7.17), and using the fact that $S(h) = S(\frac{1}{h})$, $(h > 0)$ (see [226, Lemma 2.2]), we infer $\frac{m\nabla_v M}{m\sharp_v M} \le S(\frac{M}{m})$.

As an application of Theorem 2.7.1, we prove the following order among geometric means as a tool for reversing the Golden–Thompson inequality.

Proposition 2.7.4. *For $A, B > 0$ satisfying $0 < mI \le A, B \le MI$ with $0 < m < M$, for $r = \min\{v, 1-v\}$ with $v \in [0,1]$ and for each $0 < q \le p$, there exist unitary matrices U and V such that*

$$\left(\frac{M^p !_v m^p}{M^p \sharp_r m^p}\right)^{\frac{1}{p}} V(A^p \sharp_v B^p)^{\frac{1}{p}} V^* \le (A^q \sharp_v B^q)^{\frac{1}{q}} \le \left(\frac{m^p \nabla_r M^p}{m^p \sharp_r M^p}\right)^{\frac{1}{p}} U(A^p \sharp_v B^p)^{\frac{1}{p}} U^*. \tag{2.7.18}$$

Proof. According to (2.7.16), for $0 < q \le p$, we can write

$$A^{\frac{q}{p}} \sharp_v B^{\frac{q}{p}} \le \left(\frac{m\nabla_r M}{m\sharp_r M} A\sharp_v B\right)^{\frac{q}{p}}.$$

(Actually, the above inequality holds because of the operator monotonicity of the function $f(t) = t^r$ with $t > 0$, $r \in [0,1]$). Replacing A by A^p and B by B^p to get

$$A^q \sharp_v B^q \le \left(\frac{m^p \nabla_r M^p}{m^p \sharp_r M^p} A^p \sharp_v B^p\right)^{\frac{q}{p}}. \tag{2.7.19}$$

For the case $q \ge 1$, we can write

$$(A^q \sharp_v B^q)^{\frac{1}{q}} \le \left(\frac{m^p \nabla_r M^p}{m^p \sharp_r M^p} A^p \sharp_v B^p\right)^{\frac{1}{p}}.$$

For the case $0 < q \le 1$, by the **minimax principle** (See [20, Corollary III.1.2]), there exists a subspace F of co-dimension $k-1$ such that

$$\lambda_k((A^q \sharp_v B^q)^{\frac{1}{q}}) = \max_{x\in F, \|x\|=1} \langle x, (A^q \sharp_v B^q)^{\frac{1}{q}} x\rangle = \max_{x\in F, \|x\|=1} \langle x, (A^q \sharp_v B^q) x\rangle^{\frac{1}{q}}. \tag{2.7.20}$$

Therefore,

$$\begin{aligned}
\lambda_k((A^q \sharp_v B^q)^{\frac{1}{q}}) &\le \max_{x\in F, \|x\|=1} \left(\left(\frac{m^p \nabla_r M^p}{m^p \sharp_r M^p}\right)^{\frac{q}{p}} \langle x, (A^p \sharp_v B^p)^{\frac{q}{p}} x\rangle\right)^{\frac{1}{q}} \quad \text{(by (2.7.19))} \\
&\le \left(\frac{m^p \nabla_r M^p}{m^p \sharp_r M^p}\right)^{\frac{1}{p}} \max_{x\in F, \|x\|=1} (\langle x, (A^p \sharp_v B^p)^{\frac{1}{p}} x\rangle) \quad \text{(by [97, Theorem 1.4])} \\
&\le \left(\frac{m^p \nabla_r M^p}{m^p \sharp_r M^p}\right)^{\frac{1}{p}} \lambda_k((A^p \sharp_v B^p)^{\frac{1}{p}}) \quad \text{(by (2.7.20))}.
\end{aligned}$$

By replacing A by A^{-1} and B by B^{-1} in the right-hand side of inequality (2.7.18), and then taking inverse we get the left-hand side of inequality (2.7.18). □

As a direct consequence of Proposition 2.7.4, we have the following refinement of the reverse of the **Golden–Thompson inequality** [102, 224] which refines [219, Theorem 3.4].

Corollary 2.7.5. *Let H and K be Hermitian matrices such that $mI \le H, K \le MI$ for some scalars $m \le M$, and let $v \in [0,1]$. Then for any unitarily invariant norm $\|\cdot\|_u$,*

$$\|e^{H\nabla_v K}\|_u \le \left(\frac{e^{pm}\nabla_r e^{pM}}{e^{pm}\sharp_r e^{pM}}\right)^{\frac{1}{p}} \|(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}}\|_u$$

holds for all $p > 0$. In particular,

$$\|e^{H+K}\|_u \le \left(\frac{e^{2m}\nabla_r e^{2M}}{e^{2m}\sharp_r e^{2M}}\right)\|e^{2H}\sharp e^{2K}\|_u.$$

Proof. Replacing A and B by e^H and e^K in Proposition 2.7.4, respectively, it follows that for each $0 < q \le p$ there exist unitary matrix $U_{p,q}$ such that

$$(e^{qH}\sharp_v e^{qK})^{\frac{1}{q}} \le \left(\frac{e^{pm}\nabla_r e^{pM}}{e^{pm}\sharp_r e^{pM}}\right)^{\frac{1}{p}} U_{p,q}(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}} U_{p,q}^*.$$

Since $e^{H\nabla_v K} = \lim_{q\to 0}(e^{qH}\sharp_v e^{qK})^{\frac{1}{q}}$ (see [114, Lemma 3.3]), it follows that for each $p > 0$ there exist unitary matrix U such that

$$e^{H\nabla_v K} \le \left(\frac{e^{pm}\nabla_r e^{pM}}{e^{pm}\sharp_r e^{pM}}\right)^{\frac{1}{p}} U(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}} U^*. \tag{2.7.21}$$

The desired result follows directly from (2.7.21). □

Furthermore, we have the following complementary result of Theorem 2.7.1 for operator monotone decreasing functions.

Theorem 2.7.2 ([104]). *Let $g : (0,\infty) \to (0,\infty)$ be an operator monotone decreasing function and let $A, B \in \mathbb{B}(\mathcal{H})$ be such that $mI \le A, B \le MI$ for some scalars $0 < m < M$. Then*

$$g(A\sharp_v B) \le \frac{m\nabla_r M}{m\sharp_r M}(g(A)\sharp_v g(B)),$$

where $r = \min\{v, 1-v\}$ and $v \in [0,1]$.

Proof. We have

$$\begin{aligned} g(A)\sharp_v g(B) &\ge g(A\nabla_v B) \quad \text{(by (2.16) in [8])} \\ &\ge g\left(\frac{m\nabla_r M}{m\sharp_r M}A\sharp_v B\right) \quad \text{(by (2.7.1))} \\ &\ge \frac{m\sharp_r M}{m\nabla_r M}g(A\sharp_v B) \quad \text{(by Lemma 2.7.2 (ii))}. \end{aligned} \tag{2.7.22}$$

This completes the proof. □

Our last result in this direction reads as follows.

Corollary 2.7.6. *Let $g : (0,\infty) \to (0,\infty)$ be an operator monotone decreasing function and let $A, B \geq 0$ be such that $mI \leq A^p, B^q \leq MI$ with $0 < m < M$. If $\frac{1}{p} + \frac{1}{q} = 1$, $p, q > 1$, then for all $\xi \in \mathcal{H}$,*

$$g(A^p \sharp_{\frac{1}{q}} B^q) \leq \frac{m\nabla_\alpha M}{m\sharp_\alpha M}(g(A^p)\sharp_{\frac{1}{q}} g(B^q)) \tag{2.7.23}$$

and

$$\langle g(A^p \sharp_{\frac{1}{q}} B^q)\xi, \xi\rangle \leq \frac{m\nabla_\alpha M}{m\sharp_\alpha M}\langle g(A^p)\xi, \xi\rangle^{\frac{1}{p}} \langle g(B^q)\xi, \xi\rangle^{\frac{1}{q}}, \tag{2.7.24}$$

where $r = \min\{1/p, 1/q\}$.

Proof. One can prove (2.7.23) directly using Theorem 2.7.2. We prove the inequality (2.7.24). The inequality (2.7.22) obviously ensures

$$g\left(\frac{m\nabla_r M}{m\sharp_r M} A^p \sharp_{\frac{1}{q}} B^q\right) \leq g(A^p \nabla_{\frac{1}{q}} B^q). \tag{2.7.25}$$

Therefore,

$$\begin{aligned}
\langle g(A^p \sharp_{\frac{1}{q}} B^q)\xi, \xi\rangle &\leq \frac{m\nabla_r M}{m\sharp_r M}\left\langle g\left(\frac{m\nabla_r M}{m\sharp_r M} A^p \sharp_{\frac{1}{q}} B^q\right)\xi, \xi\right\rangle \quad \text{(by Lemma 2.7.2 (b))} \\
&\leq \frac{m\nabla_r M}{m\sharp_r M}\langle g(A^p \nabla_{\frac{1}{q}} B^q)\xi, \xi\rangle \quad \text{(by (2.7.25))} \\
&\leq \frac{m\nabla_r M}{m\sharp_r M}\langle g(A^p)\sharp_{\frac{1}{q}} g(B^q)\xi, \xi\rangle \quad \text{(by (2.16) in [8])} \\
&\leq \frac{m\nabla_r M}{m\sharp_r M}\langle g(A^p)\xi, \xi\rangle^{\frac{1}{p}} \langle g(B^q)\xi, \xi\rangle^{\frac{1}{q}} \quad \text{(by [29, Lemma 8])},
\end{aligned}$$

and the proof is complete. □

2.8 Sharp inequalities for operator means

The **operator Pólya–Szegö inequality** [138, Theorem 4] is given as follows:

$$\Phi(A)\sharp\Phi(B) \leq \frac{M+m}{2\sqrt{Mm}}\Phi(A\sharp B), \tag{2.8.1}$$

whenever $mI \leq A, B \leq MI$, Φ is a positive linear map on $\mathbb{B}(\mathcal{H})$ and m, M are positive numbers. D. T. Hoa et al. [117, Theorem 2.12] proved that if Φ is a positive linear map, f is a nonzero operator monotone function on $[0,\infty)$ and $A, B \in \mathbb{B}(\mathcal{H})$ such that $0 < mI \leq A, B \leq MI$, then

$$f(\Phi(A))\tau f(\Phi(B)) \leq \frac{(M+m)^2}{4Mm} f(\Phi(A\sigma B)), \tag{2.8.2}$$

where σ, τ are two arbitrary operator means between the arithmetic mean ∇ and the harmonic mean $!$. In this section, we extend this result to the weighted means τ_v, σ_v and under the **sandwich condition** $sA \leq B \leq tA$. Our results will be natural generalizations of (2.8.2).

2.8.1 Operator means inequalities

We begin with the following new reverse of (2.0.3).

Theorem 2.8.1 ([83]). *For $A, B > 0$ satisfying $sA \leq B \leq tA$ with $0 < s \leq t$ and for any $v \in [0, 1]$, we have*

$$\frac{1}{\xi} A \nabla_v B \leq A \sharp_v B \leq \psi A !_v B, \tag{2.8.3}$$

where $\xi = \max\{\frac{(1-v)+vs}{s^v}, \frac{(1-v)+vt}{t^v}\}$ and $\psi = \max\{s^v((1-v) + \frac{v}{s}), t^v((1-v) + \frac{v}{t})\}$.

Proof. Define

$$f_v(x) = \frac{(1-v) + vx}{x^v}, \quad \text{where } 0 < s \leq x \leq t \text{ and } v \in [0, 1].$$

Direct calculations show that $f'(x) = v(1-v)(x-1)x^{-v-1}$. Since f is continuous on the interval $[s, t]$, $f_v(x) \leq \max\{f_v(s), f_v(t)\}$. Where,

$$(1-v) + vx \leq \xi x^v. \tag{2.8.4}$$

Now using the inequality (2.8.4) with $T = A^{-\frac{1}{2}} B A^{-\frac{1}{2}}$ and applying a standard functional calculus argument, we obtain the first inequality in (2.8.3).

The second one follows by applying similar arguments to the function

$$g_v(x) = x^v\left((1-v) + \frac{v}{x}\right), \quad \text{where } 0 < s \leq x \leq t \text{ and } v \in [0, 1]. \qquad \square$$

For the functions f_v and g_v defined in Theorem 2.8.1, we have the following inequalities that will be used later.

Lemma 2.8.1. *Let f_v and g_v be the functions defined in the proof of Theorem 2.8.1 for $x > 0$. Then we have the following properties:*
(i) *For $0 < x \leq 1$ and $0 \leq v \leq \frac{1}{2}$, we have $f_v(x) \leq f_v(\frac{1}{x})$ and $g_v(x) \geq g_v(\frac{1}{x})$.*
(ii) *For $0 < x \leq 1$ and $\frac{1}{2} \leq v \leq 1$, we have $f_v(x) \geq f_v(\frac{1}{x})$ and $g_v(x) \leq g_v(\frac{1}{x})$.*

Proof. We set $F_v(x) = f_v(x) - f_v(\frac{1}{x})$ and $G_v(x) = g_v(x) - g_v(\frac{1}{x})$ for $x > 0$. Then we calculate $\frac{dF_v(x)}{dx} = \frac{v(1-v)(x-1)(1-x^{2v-1})}{x^{v+1}}$ and $\frac{dG_v(x)}{dx} = \frac{v(1-v)(x-1)(x^{2v-1}-1)}{x^{v+1}}$. Then a standard calculus argument implies the four implications. $\square$

The next result is a consequence of Theorem 2.8.1 with $s = \frac{m}{M}$ and $t = \frac{M}{m}$. This implies a more familiar form of the following **means inequalities**.

Corollary 2.8.1. *For $A, B > 0$ satisfying $mI \le A, B \le MI$ with $0 < m < M$, we then have*

$$\frac{m\sharp_r M}{m\nabla_r M} A\nabla_v B \le A\sharp_v B \le \frac{m\sharp_R M}{m!_R M} A!_v B, \tag{2.8.5}$$

where $r = \min\{v, 1-v\}$, $R = \max\{v, 1-v\}$, and $v \in [0, 1]$.

We would like to emphasize that (2.8.5) is an extension of [53, Theorem 13] to the weighted means (thanks to $\frac{M\sharp m}{M!m} = \frac{M\nabla m}{M\sharp m}$). We also remark that Corollary 2.8.1 has been shown in [104]. Before proceeding further, we present the following remark about the powers of operator inequalities.

Remark 2.8.1. From Corollary 2.8.1, we have the well-known inequality

$$\Phi(A\nabla B) \le \frac{m\nabla M}{m\sharp M}\Phi(A\sharp B), \tag{2.8.6}$$

for a positive linear map Φ and $0 < mI \le A, B \le MI$.

It is well known that the mapping $t \mapsto t^2$ is not operator monotone, and hence one cannot simply square both sides of (2.8.6). As for this problem, M. Lin proposed an elegant method in [142, 143]. The technique proposed in these references was then used by several authors to present powers of operator inequalities. In particular, it is shown in [47] that one can take the p-power of (2.8.6) as follows:

$$\Phi^p(A\nabla B) \le \left(\frac{(m+M)^2}{4^{\frac{2}{p}}mM}\right)^p \Phi^p(A\sharp B), \quad p \ge 2. \tag{2.8.7}$$

When $p = 2$, this gives the same conclusion to the Lin's result. We here follow a simple approach to obtain these inequalities. For this end, we need to recall the following fact; see [92, Theorem 2, p. 204]: If $A > 0$ and $0 < mI \le B \le MI$, we have

$$B \le A \Rightarrow B^p \le \frac{(M^{p-1} + m^{p-1})^2}{4m^{p-1}M^{p-1}} A^p, \quad p \ge 2. \tag{2.8.8}$$

Now applying (2.8.8) on (2.8.6), we obtain

$$\Phi^p(A\nabla B) \le \frac{(M^{p-1} + m^{p-1})^2}{4m^{p-1}M^{p-1}} \left(\frac{M+m}{2\sqrt{mM}}\right)^p \Phi^p(A\sharp B), \quad p \ge 2. \tag{2.8.9}$$

When $p = 2$, we obtain the same squared version as **Lin's inequality**. However, for other values of p, it is interesting to compare (2.8.7) with (2.8.9). For this end, one can define the function

$$f_p(x) = \left(\frac{(x+1)^2}{4^{2/p}x}\right)^p - \frac{(1 + x^{p-1})^2}{4x^{p-1}}\left(\frac{1+x}{2\sqrt{x}}\right)^p, \quad x \ge 1, p \ge 2.$$

Direct calculations show that $f_{2.5}(7) > 0$ while $f_5(8) < 0$, which means that neither (2.8.7) nor (2.8.9) is uniformly better than the other. However, this entails the following refinement of (2.8.7) and (2.8.9):

$$\Phi^p(A\nabla B) \le \eta \Phi^p(A\sharp B),$$

where

$$\eta = \min\left\{\left(\frac{(m+M)^2}{4^{\frac{2}{p}} mM}\right)^p, \frac{(M^{p-1}+m^{p-1})^2}{4m^{p-1}M^{p-1}}\left(\frac{M+m}{2\sqrt{mM}}\right)^p\right\}.$$

In a similar manner, we can also obtain the following inequality:

$$\Phi^p(A\nabla B) \le \eta\big(\Phi(A)\sharp\Phi(B)\big)^p.$$

Proposition 2.8.1. *Let A, B be two positive operators and m_1, m_2, M_1, M_2 be positive real numbers, satisfying $0 < m_2 I \le A \le m_1 I < M_1 I \le B \le M_2 I$. Then, for $0 \le v \le 1$,*

$$\frac{m_2\sharp_v M_2}{m_2\nabla_v M_2} A\nabla_v B \le A\sharp_v B \le \frac{m_1\sharp_v M_1}{m_1\nabla_v M_1} A\nabla_v B, \tag{2.8.10}$$

and

$$\frac{m_1\sharp_v M_1}{m_1 !_v M_1} A!_v B \le A\sharp_v B \le \frac{m_2\sharp_v M_2}{m_2 !_v M_2} A!_v B. \tag{2.8.11}$$

Remark 2.8.2. From Proposition 2.8.1, we find for $v \in [0,1]$ that, if $0 < a_n \le a_{n-1} \le \cdots \le a_1 < b_1 \le \cdots \le b_{n-1} \le b_n$, then

$$1 \le \frac{a_1\nabla_v b_1}{a_1\sharp_v b_1} \le \cdots \le \frac{a_{n-1}\nabla_v b_{n-1}}{a_{n-1}\sharp_v b_{n-1}} \le \frac{a_n\nabla_v b_n}{a_n\sharp_v b_n},$$

and

$$1 \le \frac{a_1\sharp_v b_1}{a_1 !_v b_1} \le \cdots \le \frac{a_{n-1}\sharp_v b_{n-1}}{a_{n-1} !_v b_{n-1}} \le \frac{a_n\sharp_v b_n}{a_n !_v b_n}.$$

Remark 2.8.3. We comment on the sharpness of our results compared to some known results in the literature.

(i) The constants in Theorem 2.8.1, Corollary 2.8.1 and Proposition 2.8.1 are best possible in the sense that smaller quantities cannot replace them. As a matter of fact, f_v and g_v are continuous functions on $0 < s \le x \le t$, so that $f_v(x) \le f_v(t)$ is a sharp inequality for example. For the reader convenience, we will show that our results are stronger than the inequalities obtained in [140, 248].
On account of [248, Corollary 3], if $a, b > 0$ and $v \in [0,1]$, we have

$$K^r(t)a\sharp_v b \le a\nabla_v b,$$

where $r = \min\{v, 1-v\}$ and $t = \frac{b}{a}$. Letting $a = m_1$ and $b = M_1$ we get

$$\frac{m_1 \sharp_v M_1}{m_1 \nabla_v M_1} \le \frac{1}{K^r(\frac{M_1}{m_1})}.$$

In addition, W.Liao et al. in [140, Corollary 2.2] proved that

$$a \nabla_v b \le K^R(t) a \sharp_v b,$$

where $R = \max\{v, 1-v\}$. By choosing $a = m_2$ and $b = M_2$, we have

$$\frac{m_2 \nabla_v M_2}{m_2 \sharp_v M_2} \le K^R\left(\frac{M_2}{m_2}\right).$$

(ii) The assumption on A and B ($sA \le B \le tA$ in Theorem 2.8.1) is more general than $mI \le A, B \le MI$ in Corollary 2.8.1 and the conditions (i) or (ii) in Lemma 2.8.1. The conditions (i) or (ii) in Lemma 2.8.1 imply $I \le \frac{M_1}{m_1} I \le A^{-1/2} B A^{-1/2} \le \frac{M_2}{m_2} I$ which is a special case of $0 < sI \le A^{-1/2} B A^{-1/2} \le tI$ with $s = \frac{M_1}{m_1}$ and $t = \frac{M_2}{m_2}$.

An application of Proposition 2.8.1 under a positive linear map is given as follows.

Corollary 2.8.2. *Let A, B be two positive operators, Φ be a normalized positive linear map, $v \in [0,1]$ and m_1, m_2, M_1, M_2 be positive real numbers.*
(i) *If $0 < m_2 I \le A \le m_1 I \le M_1 I \le B \le M_2 I$, then*

$$\Phi(A) \sharp_v \Phi(B) \le \frac{m_1 \sharp_v M_1}{m_2 \sharp_v M_2} \frac{m_2 \nabla_v M_2}{m_1 \nabla_v M_1} \Phi(A \sharp_v B). \tag{2.8.12}$$

(ii) *If $0 < m_2 I \le B \le m_1 I \le M_1 I \le A \le M_2 I$, then*

$$\Phi(A) \sharp_v \Phi(B) \le \frac{M_1 \sharp_v m_1}{M_2 \sharp_v m_2} \frac{M_2 \nabla_v m_2}{M_1 \nabla_v m_1} \Phi(A \sharp_v B). \tag{2.8.13}$$

Remark 2.8.4. In the special case when $v = \frac{1}{2}$, our inequalities in Corollary 2.8.2 improve inequality (2.8.1). This follows from the fact that $\frac{m_1 \sharp M_1}{m_1 \nabla M_1} \le 1$ and $\frac{M_1 \sharp m_1}{M_1 \nabla m_1} \le 1$.

In the following, we present related results for $v \notin [0,1]$. It is easy to prove that if $a, b > 0$ and $v \notin [0,1]$, then $a \nabla_v b \le a \natural_v b$, which implies

$$A \nabla_v B \le A \natural_v B, \tag{2.8.14}$$

whenever $A, B \in \mathbb{B}(\mathcal{H})$ are two positive operators.

The following result provides a multiplicative refinement and reverse of inequality (2.8.14). We omit the details of the proof since it is similar to the proof of Proposition 2.8.1.

Proposition 2.8.2. *Let A, B be two positive operators and m_1, m_2, M_1, M_2 be positive real numbers. If $0 < m_2 I \le A \le m_1 I < M_1 I \le B \le M_2 I$, then*

$$\frac{m_1 \natural_v M_1}{m_1 \nabla_v M_1} A \nabla_v B \le A \natural_v B \le \frac{m_2 \natural_v M_2}{m_2 \nabla_v M_2} A \nabla_v B \quad \text{for } v > 1$$

and

$$\frac{m_1 !_v M_1}{m_1 \natural_v M_1} A \natural_v B \le A !_v B \le \frac{m_2 !_v M_2}{m_2 \natural_v M_2} A \natural_v B \quad \text{for } v < 0.$$

Next, we present operator inequalities involving a positive linear map and operator monotone functions. We begin with the following application of Theorem 2.8.1.

Theorem 2.8.2 ([83]). *Let Φ be a positive linear map, A, B be positive operators such that $sA \le B \le tA$ for some scalars $0 < s \le t$, and let $!_v \le \sigma_v, \tau_v \le \nabla_v$ for any $v \in [0, 1]$.*

If f is a nonzero operator monotone function on $[0, \infty)$, then

$$f(\Phi(A)) \tau_v f(\Phi(B)) \le \xi\psi f(\Phi(A \sigma_v B)). \tag{2.8.15}$$

If g is a nonzero operator monotone decreasing function on $[0, \infty)$, then

$$g(\Phi(A \sigma_v B)) \le \xi\psi(g(\Phi(A)) \tau_v g(\Phi(B))),$$

where $\xi = \max\{\frac{(1-v)+vs}{s^v}, \frac{(1-v)+vt}{t^v}\}$ and $\psi = \max\{s^v((1-v)+\frac{v}{s}), t^v((1-v)+\frac{v}{t})\}$.

Remark 2.8.5. Taking $v = \frac{1}{2}$, $s = \frac{m}{M}$ and $t = \frac{M}{m}$ in Theorem 2.8.2, we have $\xi = \psi = \frac{1}{2}(\sqrt{\frac{m}{M}} + \sqrt{\frac{M}{m}})$. Since $\xi\psi = \frac{(M+m)^2}{4Mm}$, the inequality (2.8.15) recovers the inequality (2.8.2).

Remark 2.8.6. In Theorem 2.8.2, it is proved that for two means τ_v, σ_v, we have

$$f(\Phi(A)) \tau_v f(\Phi(B)) \le \xi\psi f(\Phi(A \sigma_v B)), \tag{2.8.16}$$

if f is a nonzero operator monotone function on $[0, \infty)$.

We can modify the constant $\xi\psi$ as follows. Take the function $h_v(x) = (1-v+vx)(1-v+\frac{v}{x})$, $(v \in [0,1], s \le x \le t)$. Direct computations show that

$$h_v'(x) = \frac{v(1-v)(x^2-1)}{x^2},$$

which implies

$$h_v(x) \le \max\{h_v(s), h_v(t)\} := \alpha.$$

If we have a condition $!_v \le \tau_v, \sigma_v \le \nabla_v$, then we obtain by the similar way to the proof of Theorem 2.8.2,

$$f(\Phi(A)) \tau_v f(\Phi(B)) \le \alpha f(\Phi(A \sigma_v B)). \tag{2.8.17}$$

That is, the constant $\xi\psi$ has been replaced by α. Notice that $\xi\psi = \alpha$ in case both maximum values (for ξ, ψ) are attained at the same t or s. If $s, t \le 1$ or $s, t \ge 1$, we do have $\alpha = \xi\psi$. But, if $s < 1$ and $t > 1$, it can be seen that $\alpha \le \xi\psi$, which will be a better approximation.

Notice that Theorem 2.8.2 is a ratio inequality. The next result is a difference version, where upper bounds of the difference between $f(\Phi(A))\tau f(\Phi(B))$ and $f(\Phi(A\sigma B))$ are given.

Corollary 2.8.3. *Let Φ be a normalized positive linear map, A, B be two positive operators such that $mI \le A, B \le MI$ for some scalars $0 < m < M$, and let $! \le \sigma, \tau \le \nabla$.*

If f is a nonzero operator monotone function on $[0, \infty)$, then

$$f(\Phi(A))\tau f(\Phi(B)) - f(\Phi(A\sigma B)) \le \frac{(M-m)^2}{4Mm} f(M)I.$$

Further, if g is a nonzero operator monotone decreasing function on $[0, \infty)$, then

$$g(\Phi(A\sigma B)) - g(\Phi(A))\tau g(\Phi(B)) \le \frac{(M-m)^2}{4Mm} g(m)I.$$

Corollary 2.8.4. *Let A, B be as in Theorem 2.8.1 and let $g : (0, \infty) \to (0, \infty)$ be an operator monotone decreasing function. If σ_v is a symmetric mean between ∇_v and $!_v$, $0 \le v \le 1$, then for any vector $x \in \mathcal{H}$,*

$$\langle g(A\sigma_v B)x, x\rangle \le \xi\psi \langle g(A)x, x\rangle^{1-v} \langle g(B)x, x\rangle^{v},$$

for the same ξ, ψ from above.

The following lemma, which we need in our analysis, can be found in [209, Lemma 3.11].

Lemma 2.8.2. *Let $f : (0, \infty) \to (0, \infty)$ be operator convex, $A, B \in \mathbb{B}(\mathcal{H})$ be two self-adjoint operators and let $v \notin [0, 1]$. Then*

$$f(A)\nabla_v f(B) \le f(A\nabla_v B), \tag{2.8.18}$$

and the reverse inequality holds if f is operator concave.

Proposition 2.8.3. *For $A, B > 0$ and $m_1, m_2, M_1, M_2 > 0$, we have the following:*

(i) *If $0 < m_2 I \le A \le m_1 I < M_1 I \le B \le M_2 I$, then*

$$f(A\natural_v B) \le \frac{m_2 \natural_v M_2}{m_2 \nabla_v M_2}(f(A)\natural_v f(B)) \quad \text{for } v > 1 \tag{2.8.19}$$

for any operator monotone function $f : (0, \infty) \to (0, \infty)$, and

$$g(A)\natural_v g(B) \le \frac{m_2 \natural_v M_2}{m_2 \nabla_v M_2} g(A\natural_v B) \quad \text{for } v > 1 \tag{2.8.20}$$

for any operator monotone decreasing function $g : (0, \infty) \to (0, \infty)$.

(ii) *If* $0 < m_2 I \le B \le m_1 I < M_1 I \le A \le M_2 I$, *then*

$$f(A \natural_v B) \le \frac{M_2 \natural_v m_2}{M_2 \nabla_v m_2} (f(A) \natural_v f(B)) \quad \text{for } v < 0$$

for any operator monotone function $f : (0,\infty) \to (0,\infty)$, *and*

$$g(A) \natural_v g(B) \le \frac{M_2 \natural_v m_2}{M_2 \nabla_v m_2} g(A \natural_v B) \quad \text{for } v < 0$$

for any operator monotone decreasing function $g : (0,\infty) \to (0,\infty)$.

2.8.2 Young inequalities by sharp constants

We review a few refinements of Young inequality. H. Zuo et al. showed in [248, Theorem 7] that the following inequality holds:

$$K^r(h) A \sharp_v B \le A \nabla_v B, \quad r = \min\{v, 1-v\}, \quad K(h) = \frac{(h+1)^2}{4h}, \quad h = \frac{M}{m} \tag{2.8.21}$$

whenever $0 < m'I \le B \le mI < MI \le A \le M'I$ or $0 < m'I \le A \le mI < MI \le B \le M'I$. As the authors mentioned in [248], the inequality (2.8.21) improves the following refinement of Young's inequality involving the Specht ratio given in (2.0.5),

$$S(h^r) A \sharp_v B \le A \nabla_v B.$$

Under the above assumptions, S. S. Dragomir proved in [42, Corollary 1] that

$$A \nabla_v B \le \exp\left(\frac{v(1-v)}{2} (h-1)^2 \right) A \sharp_v B.$$

We remark that there is no relationship between the constants $K^r(h)$ and $\exp(\frac{v(1-v)}{2}(h-1)^2)$ in general.

In [104, 83], we proved some sharp ratio reverses of Young's inequality. In this subsection, as the continuation of this chapter, we establish sharp bounds for the arithmetic, geometric and harmonic mean inequalities. Moreover, we shall show some difference-type refinements and reverses of Young's inequality. We will formulate our new results in a more general setting, namely the sandwich assumption $sA \le B \le tA$ $(0 < s \le t)$. Additionally, we present some Young-type inequalities for the wide range of v; that is, $v \notin [0,1]$.

In the paper [83], we gave a reverse Young inequalities with new sharp constant. In this subsection, we first give a refined Young inequalities with a new sharp constant, as limited cases in the first inequalities both (i) and (ii) of the following theorem.

Theorem 2.8.3 ([80, Theorem 2.1]). *Let* $A, B > 0$ *such that* $sA \le B \le tA$ *for some scalars* $0 < s \le t$ *and let* $f_v(x) = \frac{(1-v)+vx}{x^v}$ *for* $x > 0$, *and* $v \in [0,1]$.

(i) *If* $t \le 1$, *then* $f_v(t)A\sharp_v B \le A\nabla_v B \le f_v(s)A\sharp_v B$.
(ii) *If* $s \ge 1$, *then* $f_v(s)A\sharp_v B \le A\nabla_v B \le f_v(t)A\sharp_v B$.

Remark 2.8.7. It is worth emphasizing that each assertion in Theorem 2.8.3, implies the other one. For instance, assume that the assertion (ii) holds, that is,

$$f_v(s) \le f_v(x) \le f_v(t), \quad 1 \le s \le x \le t. \tag{2.8.22}$$

Let $t \le 1$, then $1 \le \frac{1}{t} \le \frac{1}{x} \le \frac{1}{s}$. Hence (2.8.22) ensures that

$$f_v\left(\frac{1}{t}\right) \le f_v\left(\frac{1}{x}\right) \le f_v\left(\frac{1}{s}\right).$$

So

$$\frac{(1-v)t+v}{t^{1-v}} \le \frac{(1-v)x+v}{x^{1-v}} \le \frac{(1-v)s+v}{s^{1-v}}.$$

Now, by replacing v by $1-v$ we get

$$\frac{(1-v)+vt}{t^{v}} \le \frac{(1-v)+vx}{x^{v}} \le \frac{(1-v)+vs}{s^{v}}$$

which means

$$f_v(t) \le f_v(x) \le f_v(s), \quad 0 < s \le x \le t \le 1.$$

In the same spirit, we can derive (ii) from (i).

Corollary 2.8.5. *Let* $A, B > 0$, $m, m', M, M' > 0$ *and* $v \in [0,1]$.
(i) *If* $0 < m'I \le A \le mI < MI \le B \le M'I$, *then*

$$\frac{m\nabla_v M}{m\sharp_v M} A\sharp_v B \le A\nabla_v B \le \frac{m'\nabla_v M'}{m'\sharp_v M'} A\sharp_v B. \tag{2.8.23}$$

(ii) *If* $0 < m'I \le B \le mI < MI \le A \le M'I$, *then*

$$\frac{M\nabla_v m}{M\sharp_v m} A\sharp_v B \le A\nabla_v B \le \frac{M'\nabla_v m'}{M'\sharp_v m'} A\sharp_v B. \tag{2.8.24}$$

Note that the second inequalities in both (i) and (ii) of Theorem 2.8.3 and Corollary 2.8.5 are special cases of [83, Theorem A].

Remark 2.8.8. It is remarkable that the inequalities $f_v(t) \le f_v(x) \le f_v(s)$ $(0 < s \le x \le t \le 1)$ given in the proof of Theorem 2.8.3 are sharp, since the function $f_v(x)$ for $s \le x \le t$ is continuous. So, all results given from Theorem 2.8.3 are similarly sharp. As a matter of fact, let $A = MI$ and $B = mI$, then from the left-hand side of (2.8.24), we infer

$$A\nabla_v B = (M\nabla_v m)I \quad \text{and} \quad A\sharp_v B = (M\sharp_v m)I.$$

Consequently,

$$\frac{M\nabla_v m}{M\sharp_v m} A\sharp_v B = A\nabla_v B.$$

To see that the constant $\frac{m\nabla_v M}{m\sharp_v M}$ in the left-hand side of (2.8.23) cannot be improved, we consider $A = mI$ and $B = MI$, then

$$\frac{m\nabla_v M}{m\sharp_v M} A\sharp_v B = A\nabla_v B.$$

By replacing A, B by A^{-1}, B^{-1}, respectively, then the refinement and reverse of non-commutative geometric-harmonic mean inequality can be obtained as follows.

Corollary 2.8.6. *Let $A, B > 0$, $m, m', M, M' > 0$ and $v \in [0, 1]$.*
(i) *If $0 < m'I \le A \le mI < MI \le B \le M'I$, then*

$$\frac{m'!_v M'}{m'\sharp_v M'} A\sharp_v B \le A!_v B \le \frac{m!_v M}{m\sharp_v M} A\sharp_v B.$$

(ii) *If $0 < m'I \le B \le mI < MI \le A \le M'I$, then*

$$\frac{M'!_v m'}{M'\sharp_v m'} A\sharp_v B \le A!_v B \le \frac{M!_v m}{M\sharp_v m} A\sharp_v B.$$

Now, we give a **difference reverse Young inequality with sharp constants** in the following.

Theorem 2.8.4 ([80, Theorem 2.6]). *Let $A, B > 0$ such that $sA \le B \le tA$ for some scalars $0 < s \le t$, and $v \in [0, 1]$. Then*

$$A\nabla_v B - A\sharp_v B \le \max\{g_v(s), g_v(t)\}A, \tag{2.8.25}$$

where $g_v(x) = (1 - v) + vx - x^v$ for $s \le x \le t$.

Corollary 2.8.7. *Let $A, B > 0$ such that $mI \le A, B \le MI$ for some scalars $0 < m < M$. Then for $v \in [0, 1]$,*

$$A\nabla_v B - A\sharp_v B \le \xi A,$$

where $\xi = \max\{\frac{1}{M}(M\nabla_v m - M\sharp_v m), \frac{1}{m}(m\nabla_v M - m\sharp_v M)\}$.

Remark 2.8.9. We claim that if $A, B > 0$ such that $mI \le A, B \le MI$ for some scalars $0 < m < M$ with $h = \frac{M}{m}$ and $v \in [0, 1]$, then

$$A\nabla_v B - A\sharp_v B \le \max\left\{g_v(h), g_v\left(\frac{1}{h}\right)\right\}A \le L(1, h)\log S(h)A$$

holds. Indeed, we have the inequalities

$$(1-v)+vh-h^{v} \le L(1,h)\log S(h), \quad (1-v)+v\frac{1}{h}-h^{-v} \le L(1,h)\log S(h),$$

which were originally proved in [226, Lemma 3.2], thanks to $S(h)=S(\frac{1}{h})$ and $L(1,h)=L(1,\frac{1}{h})$. Therefore, our result, Theorem 2.8.4, improves the well-known result by M. Tominaga [226, Theorem 3.1],

$$A\nabla_v B - A\sharp_v B \le L(1,h)\log S(h)A.$$

Since $g_v(x)$ is convex so that we can not obtain a general result on the lower bound for $A\nabla_v B - A\sharp_v B$. However, if we impose the conditions, we can obtain new sharp inequalities for Young inequalities as a difference-type in the first inequalities both (i) and (ii) in the following proposition. (At the same time, of course, we also obtain the upper bounds straightforwardly.)

Proposition 2.8.4. *Let $A,B>0$ such that $sA \le B \le tA$ for some scalars $0<s\le t$, $v\in[0,1]$, and g_v is defined as in Theorem 2.8.4.*
(i) *If $t\le 1$, then $g_v(t)A \le A\nabla_v B - A\sharp_v B \le g_v(s)A$.*
(ii) *If $s\ge 1$, then $g_v(s)A \le A\nabla_v B - A\sharp_v B \le g_v(t)A$.*

Corollary 2.8.8. *Let $A,B>0$, $m,m',M,M'>0$, and $v\in[0,1]$.*
(i) *If $0<m'I\le A\le mI<MI\le B\le M'I$, then*

$$\frac{1}{m}(m\nabla_v M - m\sharp_v M)A \le A\nabla_v B - A\sharp_v B \le \frac{1}{m'}(m'\nabla_v M' - m'\sharp_v M')A.$$

(ii) *If $0<m'I\le B\le mI<MI\le A\le M'I$, then*

$$\frac{1}{M}(M\nabla_v m - M\sharp_v m)A \le A\nabla_v B - A\sharp_v B \le \frac{1}{M'}(M'\nabla_v m' - M'\sharp_v m')A.$$

In the following, we use the notation $\triangledown_v$ and $\natural_v$ to distinguish from the operator means ∇_v and $\sharp_v$:

$$A\triangledown_v B=(1-v)A+vB, \quad A\natural_v B = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^{v}A^{\frac{1}{2}}$$

for $v\notin[0,1]$. Notice that, since $A,B>0$, the expressions $A\triangledown_v B$ and $A\natural_v B$ are also well-defined.

Remark 2.8.10. It is known (and easy to show) that for any $A,B>0$,

$$A\triangledown_v B \le A\natural_v B, \quad \text{for } v\notin[0,1].$$

Assume $g_v(x)$ is defined as in Theorem 2.8.4. By an elementary computation, we have

$$\begin{cases} g_v'(x)>0 & \text{for } v\notin[0,1] \text{ and } 0<x\le 1\\ g_v'(x)<0 & \text{for } v\notin[0,1] \text{ and } x>1.\end{cases}$$

Now, in the same way as above we have also for any $v\notin[0,1]$:

(i) If $0 < m'I \le A \le mI \le MI \le B \le M'I$, then

$$\frac{1}{m'}(m'\natural_v M' - m'\nabla_v M')A \le A\natural_v B - A\nabla_v B \le \frac{1}{m}(m\natural_v M - m\nabla_v M)A.$$

On account of assumptions, we also infer

$$(m'\natural_v M' - m'\nabla_v M')I \le A\natural_v B - A\nabla_v B \le (m\natural_v M - m\nabla_v M)I.$$

(ii) If $0 < m'I \le B \le mI \le MI \le A \le M'I$, then

$$\frac{1}{M}(M\natural_v m - M\nabla_v m)A \le A\natural_v B - A\nabla_v B \le \frac{1}{M'}(M'\natural_v m' - M'\nabla_v m')A.$$

On account of assumptions, we also infer

$$(M\natural_v m - M\nabla_v m)I \le A\natural_v B - A\nabla_v B \le (M'\natural_v m' - M'\nabla_v m')I.$$

In addition, with the same assumption to Theorem 2.8.4 except for $v \notin [0,1]$, we have

$$\min\{g_v(s), g_v(t)\}A \le A\nabla_v B - A\natural_v B,$$

since we have $\min\{g_v(s), g_v(t)\} \le g_v(x)$ by $g_v''(x) \le 0$, for $v \notin [0,1]$.

2.9 Complement to refined Young inequality

We say that a linear map Φ is **2-positive** if whenever the 2×2 operator matrix $\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \ge 0$, then so is $\begin{pmatrix} \Phi(A) & \Phi(B) \\ \Phi(B^*) & \Phi(C) \end{pmatrix} \ge 0$. For an operator A, such that $0 < mI \le A \le MI$ and a vector $x \in \mathcal{H}$, the following inequality is called **Kantorovich inequality** [124]:

$$\langle Ax, x\rangle\langle A^{-1}x, x\rangle \le \frac{(M+m)^2}{4Mm}, \quad \text{for } \|x\| = 1.$$

2.9.1 Refinement of Kantorovich inequality

We start from the following refined Young inequality without weighted parameter.

Lemma 2.9.1. *Let $A, B > 0$ and $1 < m < M$ satisfying $mA \le B \le MA$. Then we have*

$$\left(1 + \frac{(\log m)^2}{8}\right)A\sharp B \le A\nabla B. \tag{2.9.1}$$

Proof. First, we point out that for each $a, b > 0$,

$$\left(1 + \frac{(\log b - \log a)^2}{8}\right)\sqrt{ab} \le \frac{a+b}{2}. \tag{2.9.2}$$

This inequality plays a fundamental role in this section (for more details in this direction, see [247]).

Note that if $0 < ma \le b \le Ma$ with $1 < m < M$, then by monotonicity of logarithm function we get

$$\left(1 + \frac{(\log m)^2}{8}\right)\sqrt{ab} \le \frac{a+b}{2}. \tag{2.9.3}$$

Taking $a = 1$ in the inequality (2.9.2), we have

$$\left(1 + \frac{(\log m)^2}{8}\right)\sqrt{b} \le \frac{b+1}{2}.$$

Since $mI \le A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le MI$ and $1 < m < M$, on choosing b with the positive operator $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$, we infer from inequality (2.9.3),

$$\left(1 + \frac{(\log m)^2}{8}\right)(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}} \le \frac{A^{-\frac{1}{2}}BA^{-\frac{1}{2}} + I}{2}.$$

Multiplying both side by $A^{\frac{1}{2}}$, we deduce the desired result (2.9.1) □

As we know from [57], the following inequality is equivalent to the **Kantorovich inequality**:

$$\langle Ax, x\rangle\langle Bx, x\rangle \le \frac{(M+m)^2}{4Mm}\langle A\sharp Bx, x\rangle^2, \tag{2.9.4}$$

where $0 < mI \le A, B \le MI$ and $x \in \mathcal{H}$.

With Lemma 2.9.1 in hand, we are ready to provide a refinement of the inequality (2.9.4).

Theorem 2.9.1 ([181]). *Let $A, B \in \mathbb{B}(\mathcal{H})$ such that $0 < mI \le m'A \le B \le MI$ and $1 < m'$. Then for every unit vector $x \in \mathcal{H}$,*

$$\langle Ax, x\rangle\langle Bx, x\rangle \le \frac{(M+m)^2}{4Mm\left(1 + \frac{(\log m')^2}{8}\right)^2}\langle A\sharp Bx, x\rangle. \tag{2.9.5}$$

Proof. According to the condition $0 < mI \le m'A \le B \le MI$, we can get

$$\frac{mm'}{M}I \le A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le \frac{Mm'}{m}I.$$

It follows from the above inequality that

$$\left((A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}} - \sqrt{\frac{mm'}{M}}I\right)\left(\sqrt{\frac{Mm'}{m}}I - (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}}\right) \ge 0,$$

and easy computations yields

$$\left(\frac{(M+m)\sqrt{m'}}{\sqrt{Mm}}\right)\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^{\frac{1}{2}} \geq m'I + A^{-\frac{1}{2}}BA^{-\frac{1}{2}}. \tag{2.9.6}$$

Multiplying both sides by $A^{\frac{1}{2}}$ to inequality (2.9.6), we obtain

$$\left(\frac{(M+m)\sqrt{m'}}{\sqrt{Mm}}\right)A\sharp B \geq m'A + B.$$

Hence for every unit vector $x \in \mathcal{H}$ we have

$$\left(\frac{(M+m)\sqrt{m'}}{\sqrt{Mm}}\right)\langle A\sharp Bx, x\rangle \geq m'\langle Ax, x\rangle + \langle Bx, x\rangle.$$

Now, by using (2.9.3) for above inequality we can find that

$$\left(\frac{(M+m)\sqrt{m'}}{\sqrt{Mm}}\right)\langle A\sharp Bx, x\rangle \geq m'\langle Ax, x\rangle + \langle Bx, x\rangle \geq 2\left(1 + \frac{(\log m')^2}{8}\right)\sqrt{m'\langle Ax, x\rangle\langle Bx, x\rangle}.$$

Square both sides, we obtain the desired result (2.9.5). □

Remark 2.9.1. If we choose $B = A^{-1}$, we get from Theorem 2.9.1 that

$$\langle Ax, x\rangle\langle A^{-1}x, x\rangle \leq \frac{(M+m)^2}{4Mm\left(1 + \frac{(\log m')^2}{8}\right)^2}, \tag{2.9.7}$$

for each $x \in \mathcal{H}$ with $\|x\| = 1$.

In this case the relation (2.9.7) represents the **refined Kantorovich inequality**.

The following **reverse of the Hölder–McCarthy inequality** is well known and easily proved using Kantorovich inequality:

$$\langle A^2x, x\rangle \leq \frac{(M+m)^2}{4Mm}\langle Ax, x\rangle^2, \quad \text{for } \|x\| = 1. \tag{2.9.8}$$

Applying inequality (2.9.7), we get the following corollary that is a refinement of (2.9.8). It can be proven by the similar method in [97, Theorem 1.29].

Corollary 2.9.1. *Substituting* $\frac{A^{\frac{1}{2}}x}{\|A^{\frac{1}{2}}x\|}$ *for a unit vector* x *in Remark* 2.9.1, *we have*

$$\frac{\langle AA^{\frac{1}{2}}x, A^{\frac{1}{2}}x\rangle}{\|A^{\frac{1}{2}}x\|^2}\frac{\langle A^{-1}A^{\frac{1}{2}}x, A^{\frac{1}{2}}x\rangle}{\|A^{\frac{1}{2}}x\|^2} \leq \frac{(M+m)^2}{4Mm\left(1 + \frac{(\log m')^2}{8}\right)^2},$$

which is equivalent to saying that

$$\langle A^2x, x\rangle \leq \frac{(M+m)^2}{4Mm\left(1 + \frac{(\log m')^2}{8}\right)^2}\langle Ax, x\rangle^2, \tag{2.9.9}$$

for each $x \in \mathcal{H}$ *with* $\|x\| = 1$.

A discussion of order-preserving properties of increasing functions through the Kantorovich inequality is presented by M. Fujii, S. Izumino, R. Nakamoto and Y. Seo in [57]. They showed that if $A, B > 0$, $B \geq A$ and $0 < mI \leq A \leq MI$, then

$$A^2 \leq \frac{(M+m)^2}{4Mm} B^2. \tag{2.9.10}$$

The following result provides an improvement of inequality (2.9.10).

Proposition 2.9.1. *Let* $A, B \in \mathbb{B}(\mathcal{H})$ *such that* $0 < mI \leq m'A \leq A^{-1} \leq MI$, $1 < m'$ *and* $A \leq B$. *Then*

$$A^2 \leq \frac{(M+m)^2}{4Mm\left(1 + \frac{(\log m')^2}{8}\right)^2} B^2, \tag{2.9.11}$$

Proof. For each $x \in \mathcal{H}$ with $\|x\| = 1$, we have

$$\begin{aligned}
\langle A^2 x, x\rangle &\leq \frac{(M+m)^2}{4Mm\left(1 + \frac{(\log m')^2}{8}\right)^2} \langle Ax, x\rangle^2 \quad \text{(by (2.9.9))} \\
&\leq \frac{(M+m)^2}{4Mm\left(1 + \frac{(\log m')^2}{8}\right)^2} \langle Bx, x\rangle^2 \quad \text{(since } A \leq B\text{)} \\
&\leq \frac{(M+m)^2}{4Mm\left(1 + \frac{(\log m')^2}{8}\right)^2} \langle B^2 x, x\rangle \quad \text{(by the Hölder–McCarthy inequality),}
\end{aligned}$$

as desired. □

In [189], using the operator geometric mean, R. Nakamoto and N. Nakamura proved that

$$\Phi(A) \sharp \Phi(A^{-1}) \leq \frac{M+m}{2\sqrt{Mm}}, \tag{2.9.12}$$

whenever $0 < mI \leq A \leq MI$ and Φ is a normalized positive linear map on $\mathbb{B}(\mathcal{H})$.

It is notable that, a more general case of (2.9.12) has been studied by M. S. Moslehian et al. in [187, Theorem 2.1] which is called the operator Pólya–Szegö inequality. The operator Pólya–Szegö inequality states that: Let Φ be a positive linear map. If $0 < mI \leq A, B \leq MI$ for some positive real numbers $m < M$, then

$$\Phi(A) \sharp \Phi(B) \leq \frac{M+m}{2\sqrt{Mm}} \Phi(A \sharp B). \tag{2.9.13}$$

Our second main result in this section, which is related to inequality (2.9.13) can be stated as follows.

Theorem 2.9.2 ([181]). *Let Φ be a normalized positive linear map on $\mathbb{B}(\mathcal{H})$ and let $A, B \in \mathbb{B}(\mathcal{H})$ such that $0 < mI \le m'A \le B \le MI$ and $1 < m'$. Then*

$$\Phi(A)\sharp\Phi(B) \le \frac{M+m}{2\sqrt{Mm}\left(1+\frac{(\log m')^2}{8}\right)}\Phi(A\sharp B). \tag{2.9.14}$$

Proof. According to the hypothesis, we get the order relation

$$\left(\frac{\sqrt{m'}(M+m)}{\sqrt{Mm}}\right)\Phi(A\sharp B) \ge m'\Phi(A) + \Phi(B).$$

By using Lemma 2.9.1, we get

$$\left(\frac{\sqrt{m'}(M+m)}{\sqrt{Mm}}\right)\Phi(A\sharp B) \ge m'\Phi(A) + \Phi(B) \ge 2\sqrt{m'}\left(1+\frac{(\log m')^2}{8}\right)\Phi(A)\sharp\Phi(B).$$

Rearranging terms gives the inequality (2.9.14). □

Remark 2.9.2. If we choose $B = A^{-1}$, we get from Theorem 2.9.2 that

$$\Phi(A)\sharp\Phi(A^{-1}) \le \frac{M+m}{2\sqrt{Mm}\left(1+\frac{(\log m')^2}{8}\right)}. \tag{2.9.15}$$

This is a refinement of inequality (2.9.12).

A particular case of the inequality (2.9.15) has been known for many years: Let U_j be contraction with $\sum_{j=1}^{k} U_j^* U_j = 1_{\mathcal{H}}$ $(j = 1, 2, \dots, k)$. If A is a positive operator on $\mathcal{H}$ satisfying $0 < mI \le A \le MI$ for some scalars $m < M$, then

$$\left(\sum_{j=1}^{k} U_j^* A U_j\right)\sharp\left(\sum_{j=1}^{k} U_j^* A^{-1} U_j\right) \le \frac{M+m}{2\sqrt{Mm}}.$$

This inequality, proved by B. Mond and J. Pečarič [170], reduces to the Kantorovich inequality when $k = 1$.

Corollary 2.9.2. *By* (2.9.14),

$$\left(\sum_{j=1}^{n} U_j^* A U_j\right)\sharp\left(\sum_{j=1}^{n} U_j^* A^{-1} U_j\right) \le \frac{M+m}{2\sqrt{Mm}\left(1+\frac{(\log m')^2}{8}\right)}. \tag{2.9.16}$$

Inequality (2.9.16) *follows quite simply by noting that* $\Phi(A) = \sum_{j=1}^{k} U_j^* A U_j$ *defines a normalized positive linear map on* $\mathbb{B}(\mathcal{H})$.

2.9.2 Operator inequalities for positive linear maps

We know that if Φ is a normalized positive linear map on $\mathbb{B}(\mathcal{H})$ and $A, B \in \mathbb{B}(\mathcal{H})$ such that $0 < mI \le A, B \le MI$, then

$$\Phi^2(A\nabla B) \le K^2(h)\Phi^2(A\sharp B), \tag{2.9.17}$$

and

$$\Phi^2(A\nabla B) \le K^2(h)\big(\Phi(A)\sharp\Phi(B)\big)^2, \tag{2.9.18}$$

hold, where Φ is a normalized positive linear map and $h = \frac{M}{m}$.

In this section, we are devoted to obtain a better bound than (2.9.17) and (2.9.18).

Theorem 2.9.3 ([181]). *Let A and B be two positive operators such that $0 < mI \le A \le m'I \le M'I \le B \le MI$. Then for a normalized positive linear map Φ,*

$$\Phi^2(A\nabla B) \le \frac{K^2(h)}{\left(1 + \frac{\left(\log \frac{M'}{m'}\right)^2}{8}\right)^2}\Phi^2(A\sharp B), \tag{2.9.19}$$

and

$$\Phi^2(A\nabla B) \le \frac{K^2(h)}{\left(1 + \frac{\left(\log \frac{M'}{m'}\right)^2}{8}\right)^2}\big(\Phi(A)\sharp\Phi(B)\big)^2, \tag{2.9.20}$$

where $h = \frac{M}{m}$.

Proof. We intend to prove

$$A\nabla B + Mm\left(1 + \frac{\left(\log \frac{M'}{m'}\right)^2}{8}\right)(A\sharp B)^{-1} \le (M+m)I. \tag{2.9.21}$$

According to the hypothesis, we have

$$\frac{1}{2}(MI - A)(mI - A)A^{-1} \le 0,$$

by easy computation we find that

$$\frac{A}{2} + Mm\frac{A^{-1}}{2} \le \left(\frac{M+m}{2}\right)I \tag{2.9.22}$$

and similar argument shows that

$$\frac{B}{2} + Mm\frac{B^{-1}}{2} \le \left(\frac{M+m}{2}\right)I. \tag{2.9.23}$$

Summing up (2.9.22) and (2.9.23), we get

$$A\nabla B + Mm\frac{A^{-1}+B^{-1}}{2} \le (M+m)I. \tag{2.9.24}$$

On the other hand,

$$\begin{aligned} A\nabla B + Mm\left(1+\frac{(\log\frac{M'}{m'})^2}{8}\right)(A\sharp B)^{-1} &= A\nabla B + Mm\left(1+\frac{(\log\frac{M'}{m'})^2}{8}\right)(A^{-1}\sharp B^{-1}) \\ &\le A\nabla B + Mm\frac{A^{-1}+B^{-1}}{2} \quad (\text{by } (2.9.1)) \\ &\le (M+m)I \quad (\text{by } (2.9.24)). \end{aligned}$$

Therefore, the inequality (2.9.21) is established.

Now we try to prove (2.9.19) by using the above inequality. It is not hard to see that inequality (2.9.19) is equivalent with

$$\left\|\Phi(A\nabla B)\Phi^{-1}(A\sharp B)\right\| \le \frac{(M+m)^2}{4Mm\left(1+\frac{(\log\frac{M'}{m'})^2}{8}\right)}. \tag{2.9.25}$$

In order to prove (2.9.25), we need to show

$$\Phi(A\nabla B) + Mm\left(1+\frac{(\log\frac{M'}{m'})^2}{8}\right)\Phi^{-1}(A\sharp B) \le (M+m)I, \tag{2.9.26}$$

by Lemma 2.6.2. Besides, from the Choi inequality given in Lemma 2.6.3, we prove the much stronger statement (2.9.26), that is,

$$\Phi(A\nabla B) + Mm\left(1+\frac{(\log\frac{M'}{m'})^2}{8}\right)\Phi\big((A\sharp B)^{-1}\big) \le (M+m)I. \tag{2.9.27}$$

Using linearity of Φ and inequality (2.9.21), we can easily obtain the desired result (2.9.19). The inequality (2.9.20) can be proved analogously. □

As is known, the **Wielandt inequality** [118, p. 443] states that if $0 < mI \le A \le MI$, and $x, y \in \mathcal{H}$ with $x \perp y$, then

$$\left|\langle x, Ay\rangle\right|^2 \le \left(\frac{M-m}{M+m}\right)^2 \langle x, Ay\rangle\langle y, Ay\rangle.$$

In [24], R. Bhatia and C. Davis proved an **operator Wielandt inequality** which states that if $0 < mI \le A \le MI$ and X, Y are two partial isometries on $\mathcal{H}$ whose final spaces are orthogonal to each other, then for every 2-positive linear map Φ on $\mathbb{B}(\mathcal{H})$,

$$\Phi(X^*AY)\Phi^{-1}(Y^*AY)\Phi^{-1}(Y^*AX) \le \left(\frac{M-m}{M+m}\right)^2 \Phi(X^*AX). \tag{2.9.28}$$

M. Lin [143, Conjecture 3.4], conjectured that the following assertion could be true:

$$\|\Phi(X^*AY)\Phi^{-1}(Y^*AY)\Phi^{-1}(Y^*AX)\Phi^{-1}(X^*AX)\| \le \left(\frac{M-m}{M+m}\right)^2. \tag{2.9.29}$$

X. Fu and C. He [47] attempt to solve the conjecture and get a step closer to the conjecture. But I. H. Gümüş [103] obtained a better upper bound to approximate the right-hand side of (2.9.29) based on

$$\|\Phi(X^*AY)\Phi^{-1}(Y^*AY)\Phi^{-1}(Y^*AX)\Phi^{-1}(X^*AX)\| \le \frac{(M-m)^2}{2\sqrt{Mm}(M+m)}. \tag{2.9.30}$$

The remainder of this subsection presents an improvement for the operator version of Wielandt inequality.

Theorem 2.9.4 ([181]). *Let $0 < mI \le m'A^{-1} \le A \le MI$ and $1 < m'$ and let X and Y be two isometries such that $X^*Y = 0$. For every 2-positive linear map Φ,*

$$\|\Phi(X^*AY)\Phi^{-1}(Y^*AY)\Phi^{-1}(Y^*AX)\Phi^{-1}(X^*AX)\| \le \frac{(M-m)^2}{2\sqrt{Mm}(M+m)\left(1+\frac{(\log m')^2}{8}\right)}.$$

Proof. From (2.9.28), we have that

$$\left(\frac{M+m}{M-m}\right)^2 \Phi(X^*AY)\Phi^{-1}(Y^*AY)\Phi^{-1}(Y^*AX) \le \Phi(X^*AX).$$

Under given assumptions, we have $mI \le \Phi(X^*AX) \le MI$. Hence Proposition 2.9.1 implies

$$\left(\Phi(X^*AY)\Phi^{-1}(Y^*AY)\Phi^{-1}(Y^*AX)\right)^2 \le \frac{(M-m)^4}{4Mm(M+m)^2\left(1+\frac{(\log m')^2}{8}\right)^2}\Phi^2(X^*AX).$$

Therefore,

$$\|\Phi(X^*AY)\Phi^{-1}(Y^*AY)\Phi^{-1}(Y^*AX)\Phi^{-1}(X^*AX)\| \le \frac{(M-m)^2}{2\sqrt{Mm}(M+m)\left(1+\frac{(\log m')^2}{8}\right)},$$

which completes the proof. □

Based on inequality given in Theorem 2.9.4, we obtain a refinement of inequality (2.9.30).

2.10 On constants appearing in refined Young inequality

In this section, we show the refined Young inequality with Specht ratio and a new property of Specht ratio. Then these give an alternative proof of the refined Young

inequality with Specht ratio given in (2.2.1):

$$\frac{(1-v)+vx}{x^v} \geq S(x^r), \quad x > 0. \tag{2.10.1}$$

Finally, we consider the general property for which we showed the property on Specht ratio and Kantorovich constant. Then we state a sufficient condition that such a general property holds for a general function.

Theorem 2.10.1 ([72]). *For $v \in [0,1]$ and $x > 0$,*

$$\frac{(1-v)+vx}{x^v} \geq S^r(x). \tag{2.10.2}$$

The proof of Theorem 2.10.1 can be done by the use of Lemma 2.10.1 and Lemma 2.10.2 below with elementary (but slightly lengthy) calculations. See [72] for the details.

Lemma 2.10.1. *For $0 < x \leq 1$, we have*

$$(x-1)^2 + x(\log x)^2 \leq 2x(x-1)\log x. \tag{2.10.3}$$

Lemma 2.10.2. *For $x \geq 1$, we have*

$$\frac{\log x}{x-1} - \log\frac{\log x}{x-1} \leq \frac{2}{x+1} - \log\frac{2}{x+1}.$$

Remark 2.10.1. Since Kantorovich constant is grater than or equal to Specht ratio, we have $K^r(x) \geq S^r(x)$ for $0 \leq r \leq 1$ so that we easily obtain the inequality

$$\frac{(1-v)+xv}{x^v} \geq K^r(x) \geq S^r(x), \quad 0 \leq r \leq \frac{1}{2}$$

from the result in [248]. We also have $K(x^r) \geq S(x^r)$ for $x > 0$ and $0 \leq r \leq 1$. For the convenience of the readers, we give a proof of the inequality $K(x) \geq S(x)$ for $x > 0$, which is equivalent to $w_1(x) \geq 0$, where $w_1(x) = 2\log(x+1) - 2\log 2 - \log x - \frac{\log x}{x-1} + 1 + \log(\frac{\log x}{x-1})$. Since $\frac{\log t}{t-1} \geq \frac{2}{t+1}$ for $t > 0$, we have only to prove the inequality $w_2(x) \geq 0$, where $w_2(x) = 2\log(x+1) - 2\log 2 - \log x - \frac{\log x}{x-1} + 1 + \log(\frac{2}{x+1}) = 1 - \frac{\log x}{x-1} + \log(\frac{x+1}{2x})$. Then we have $w_2'(x) = \frac{2(1-x)+(x+1)\log x}{(x+1)(x-1)^2}$. We easily find $w_2'(x) < 0$ for $0 < x < 1$ and $w_2'(x) > 0$ for $x > 1$, since we use again the relation $\frac{\log t}{t-1} \geq \frac{2}{t+1}$ for $t > 0$.

Remark 2.10.2. For $x > 0$ and $0 \leq r \leq 1$, we have the inequality

$$K^r(x) \geq K(x^r).$$

The proof is done by the following. The above inequality is equivalent to the inequality $g_r(x) \geq 0$, where

$$g_r(x) = r\log(x+1) - \log(x^r+1) - r\log 2 + \log 2.$$

Since $g_r'(x) = \frac{r(1-x^{1-r})}{(x+1)(x^r+1)}$, we have $g_r'(1) = 0$, $g_r'(x) \leq 0$ for $0 < x \leq 1$ and $g_r'(x) \geq 0$ for $x \geq 1$. Thus we have $g_r(x) \geq g_r(1) = 0$.

Remark 2.10.3. We have no ordering between $S^r(x)$ and $K(x^r)$. See [72] for the details.

Theorem 2.10.2 ([72]). *We have the inequality*

$$S^r(x) \geq S(x^r)$$

for $x > 0$ *and* $0 \leq r \leq 1$.

The proof of Theorem 2.10.2 can be done by the use of Lemma 2.10.3 with elementary calculations. See [72] for the details.

Lemma 2.10.3. *Put* $k(t) = -(t-1)^3 + 3t(t-1)(\log t)^2 - t(t+1)(\log t)^3$. *Then* $k(t) \geq 0$ *for* $0 < t \leq 1$ *and* $k(t) \leq 0$ *for* $t \geq 1$.

Thus Theorem 2.10.1 and Theorem 2.10.2 give an alternative proof of the inequality given in (2.10.1). In addition, the inequality in (2.10.2) gives a better bound than one in (2.10.1).

It is known that Kantorovich constant and Specht ratio have similar properties:
(i) They take infinity in the limits as $x \to 0$ and $x \to \infty$.
(ii) They take 1 when $x = 1$.
(iii) They are monotone decreasing functions on $(0, 1)$.
(iv) They are monotone increasing functions on $(1, \infty)$.

As we have seen in Chapter 2, properties $K^r(x) \geq K(x^r)$ and $S^r(x) \geq S(x^r)$ for $x > 0$ and $0 \leq r \leq 1$, make us to consider the following general result. To state our proposition, we define a **geometrically convex** function $f : J \to (0, \infty)$ where $J \subset (0, \infty)$. If the function $f : J \to (0, \infty)$ where $J \subset (0, \infty)$ satisfies the inequality $f(a^{1-v}b^v) \leq f^{1-v}(a)f^v(b)$ for $a, b > 0$ and $v \in [0, 1]$, then the function f is called a geometrically convex (multiplicatively convex) function. See [192], for example, for the geometrically convex (multiplicatively convex) function. We easily find that if the geometrically convex function $f : J \to (0, \infty)$ is decreasing, then f is also (usual) convex function.

It is known that the function $f(x)$ is a geometrically convex if and only if the function $\log f(e^x)$ is (usual) convex; see for example, [14].

Proposition 2.10.1 ([72]). *Assume the function* $f : (0, \infty) \to (0, \infty)$ *satisfies* $f(1) \leq 1$. *If* f *is a geometrically convex function, then we have* $f^r(x) \geq f(x^r)$ *for* $x > 0$ *and* $0 \leq r \leq 1$.

Proof. Put $a = 1$, $b = x$ and $v = r$ in the definition of a geometrically convex function, $f(a^{1-v}b^v) \leq f^{1-v}(a)f^v(b)$. Then we obtain the inequality $f^r(x) \geq f(x^r)$ for $x > 0$ and $0 \leq r \leq 1$, since we have the condition $0 < f(1) \leq 1$. □

Remark 2.10.4. If we assume f is a twice differentiable function, we have the following alternative proof for Proposition 2.10.1. We put $g(r, x) = r\log(f(x)) - \log(f(x^r))$. It is known in [192, Proposition 4.3] that the geometrically convexity (multiplicatively convexity) of f is equivalent to the following condition (2.10.4), under the assumption

that f is twice differentiable.

$$D(x) := f(x)f'(x) + x(f(x)f''(x) - f'(x)^2) \geq 0, \quad (x > 0). \tag{2.10.4}$$

Then we calculate

$$\frac{d^2 g(r,x)}{dr^2} = -\frac{x^r(\log x)^2}{f(x^r)^2}\{f(x^r)(f'(x^r) + x^r f''(x^r)) - x^r f'(x^r)^2\} \leq 0,$$

by the condition (2.10.4). It is easy to check $g(0,x) \geq 0$, $g(1,x) = 0$ so that we have $g(r,x) \geq 0$ for any $x > 0$ and $0 \leq r \leq 1$. We thus obtain the inequality $f^r(x) \geq f(x^r)$ for $x > 0$ and $0 \leq r \leq 1$.

Thus we can make a judgment whether the function f is a geometrically convex or not, by using $D(x)$ or $\frac{d^2 \log f(e^x)}{dx^2}$. We can find $D(x) = \frac{(x+1)^2}{8x^2} \geq 0$ for Kantorovich constant $K(x)$. The expression of $D(x)$ is complicated for Specht ratio $S(x)$. But we could check that it takes nonnegative values for $x > 0$ by the numerical computations. In addition, for the function $\arg_v(x) = \frac{(1-v)+vx}{x^v}$ defined for $x > 0$ with $0 \leq v \leq 1$, we find that $D(x) = v(1-v)x^{-2v} \geq 0$ so that $\arg^r(x) \geq \arg(x^r)$ for $0 \leq r \leq 1$ and $x > 0$.

Remark 2.10.5. As for the converse in Proposition 2.10.1, under the assumption we do not impose the condition $0 < f(1) \leq 1$, the claim that the inequality $f^r(x) \geq f(x^r)$ for $x > 0, 0 \leq r \leq 1$, implies f is a geometrically convex function, is not true in general. The function $f(x) := \frac{1-x}{1+x}$ on $(0,1)$ is not a geometrically convex since $D(x) = -\frac{2(x^2+1)}{(x+1)^4} < 0$ (also $\frac{d^2 \log f(e^x)}{dx^2} = -\frac{-2e^x(e^{2x}+1)}{(e^{2x}-1)^2} < 0$). Actually, the function $f(x)$ is a geometrically concave. On the other hand, we have the inequality $f^r(x) \geq f(x^r)$ for $0 < x < 1$ and $0 \leq r \leq 1$. Indeed, we have

$$\left(\frac{1-x}{1+x}\right)^r \geq \frac{1-x}{1+x} \geq \frac{1-x^r}{1+x^r}. \tag{2.10.5}$$

The first inequality is due to $0 < \frac{1-x}{1+x} < 1$ and the second one is easy calculation as $(1-x)(1+x^r) - (1+x)(1-x^r) = 2(x^r - x) \geq 0$.

However, we have not found any examples such that the function f satisfies $0 < f(1) \leq 1$ and the inequality $f^r(x) \geq f(x^r)$, but it is not a geometrically convex. In the above example $f(x) = \frac{1-x}{1+x}$ on $(0,1)$, we see actually $f(1) := \lim_{x \nearrow 1} f(x) = 0$. We considered the open interval $(0,1)$ not $(0,1]$ in the above example, to avoid the case 0^0 in the left-hand side of the inequalities (2.10.5). We also have not found the proof that the inequality $f^r(x) \geq f(x^r)$ for the function $f : (0,\infty) \to (0,\infty)$ under the assumption $0 < f(1) \leq 1$, implies the geometrically convexity of f.

We conclude this section by stating that a property $f^r(x) \geq f(x^r)$ has been studied in [230, 231] in details by the different approach. (See also [28].) Such a property is there named as power monotonicity and applied to some known results in the framework of operator theory.

3 Inequalities related to means

In Chapter 2, we studied the developments on the refinements and reverses for Young inequality. In this chapter, we study the inequalities related to means, except for Young inequality. In Section 3.1, we give the relations for three means, the arithmetic mean $A\nabla_v B$, the geometric mean $A\sharp_v B$ and the harmonic mean $A!_v B$. In Section 3.2, we give an estimation for the Heron mean, which is defined as a linear combination between the arithmetic mean $A\nabla_v B$ and the geometric mean $A\sharp_v B$. In Section 3.3, some inequalities related to the Heron mean will be given. Finally, in Section 3.4, we give some inequalities for positive linear maps. These inequalities relate to the operator means and the famous Ando inequality.

Throughout this book, we often write the condition such as $mI \le A, B \le MI$ by $m \le A, B \le M$ for simplicity without an identity I.

3.1 Operator inequalities for three means

We start from the following proposition which gives the relations among arithmetic mean, geometric mean and harmonic mean.

Proposition 3.1.1. *For $A, B > 0$ and $r \in \mathbb{R}$, we have the following inequalities:*
(i) *If $r \ge 2$, then $rA\sharp B + (1-r)A\nabla B \le A!B$.*
(ii) *If $r \le 1$, then $rA\sharp B + (1-r)A\nabla B \ge A!B$.*

Proof. In general, by using the notion of the **representing function** $f_\sigma(x) = 1\sigma x$ for operator mean σ, it is well known [137] that $f_{\sigma_1}(x) \le f_{\sigma_2}(x)$ holds for $x > 0$ if and only if $A\sigma_1 B \le A\sigma_2 B$ holds for all positive operators A and B. Thus we can prove this proposition from the following scalar inequalities for $t > 0$. Such procedures will be used throughout this book as **Kubo–Ando theory**.
(i) $r\sqrt{t} + (1-r)\frac{t+1}{2} \le \frac{2t}{t+1}, \quad (r \ge 2).$
(ii) $r\sqrt{t} + (1-r)\frac{t+1}{2} \ge \frac{2t}{t+1}, \quad (r \le 1).$

Actually (i) above can be proven in the following way. We set $h_r(t) = \frac{2t}{t+1} - r\sqrt{t} - (1-r)\frac{t+1}{2}$, then $\frac{dh_r(t)}{dr} = -\sqrt{t} + \frac{t+1}{2} \ge 0$ implies $h_r(t) \ge h_2(t)$ for $r \ge 2$. From the relation $\frac{2t}{t+1} + \frac{t+1}{2} \ge 2\sqrt{t}$, we have $h_2(t) \ge 0$. We also give the proof for (ii) above. We set $g_r(t) = r\sqrt{t} + (1-r)\frac{t+1}{2} - \frac{2t}{t+1}$, then $\frac{dg_r(t)}{dr} = \sqrt{t} - \frac{t+1}{2} \le 0$ implies $g_r(t) \ge g_1(t)$ for $r \le 1$. From the relation $\frac{2t}{t+1} \le \sqrt{t}$, we have $g_1(t) \ge 0$. □

Remark 3.1.1. We have counterexamples of both inequalities (i) and (ii) in Proposition 3.1.1 for $1 < r < 2$. See [66] for the numerical examples.

Proposition 3.1.1 can be generalized by means of weighted parameter $v \in [0, 1]$, as the second inequality in (3.1.1) below.

https://doi.org/10.1515/9783110643473-003

Theorem 3.1.1 ([66]). *If either (i)* $0 \le v \le 1/2$ *and* $0 < A \le B$ *or (ii)* $1/2 \le v \le 1$ *and* $0 < B \le A$*, then the following inequalities hold:*

$$A\sharp B + \left(v - \frac{1}{2}\right)(B - A) \le A\sharp_v B \le \frac{1}{2}A\nabla_v B + \frac{1}{2}A!_v B. \tag{3.1.1}$$

Under the same conditions as in Theorem 3.1.1, we have $A\sharp B \ge A\sharp_v B$. In order to prove Theorem 3.1.1, it is sufficient to prove the corresponding scalar inequalities, by the procedure of the Kubo–Ando theory, shown in the proof of Proposition 3.1.1.

Lemma 3.1.1. *If either (i)* $0 \le v \le 1/2$ *and* $t \ge 1$ *or (ii)* $1/2 \le v \le 1$ *and* $0 < t \le 1$*, then the following inequalities hold:*

$$2\sqrt{t} + (2v - 1)(t - 1) \le 2t^v \le (1 - v) + vt + \left\{(1 - v) + \frac{v}{t}\right\}^{-1}. \tag{3.1.2}$$

We omit the proof of Lemma 3.1.1. See [66] for the proof of Lemma 3.1.1. We also obtain the following lemma by $\frac{\partial k_{r,v}(t)}{\partial r} = t^v - \{(1 - v) + vt\} \le 0$, for $v \in [0, 1]$ and $t > 0$.

Lemma 3.1.2. *Let* $r \in \mathbb{R}$ *and* $v \in [0, 1]$*. Then the function* $k_{r,v}(t) = rt^v + (1 - r)\{(1 - v) + vt\}$ *defined on* $t > 0$ *is monotone decreasing with respect to* r*. Therefore,* $k_{r,v}(t) \le k_{2,v}(t)$ *for* $r \ge 2$ *and* $k_{r,v}(t) \ge k_{1,v}(t)$ *for* $r \le 1$.

Lemma 3.1.2 provides the following results. Its proof follows directly from Lemma 3.1.1 and Lemma 3.1.2.

Lemma 3.1.3. *Let* $r \ge 2$*. If either (i)* $0 \le v \le 1/2$ *and* $t \ge 1$ *or (ii)* $1/2 \le v \le 1$ *and* $0 < t \le 1$*, then*

$$rt^v + (1 - r)\{(1 - v) + vt\} \le \left\{(1 - v) + \frac{v}{t}\right\}^{-1}.$$

We also have the following lemma.

Lemma 3.1.4. *Let* $r \le 1$*. For* $0 < v \le 1$ *and* $t > 0$*, we have*

$$rt^v + (1 - r)\{(1 - v) + vt\} \ge \left\{(1 - v) + \frac{v}{t}\right\}^{-1}.$$

Proof. For $r \le 1$, it follows from Lemma 3.1.2 that $rt^v + (1 - r)\{(1 - v) + vt\} \ge t^v$. Since we have $t^v \ge \{(1 - v) + \frac{v}{t}\}^{-1}$, the proof is done. □

Finally, we have the following corollary by applying Lemma 3.1.3, Lemma 3.1.4 and Theorem 3.1.1.

Corollary 3.1.1. *Let* $r \ge 2$*. If either (i)* $0 < v \le 1/2$ *and* $0 < A \le B$ *or (ii)* $1/2 \le v \le 1$ *and* $0 < B \le A$*, then*

$$rA\sharp_v B + (1 - r)A\nabla_v B \le A!_v B.$$

Let $r \le 1$*. For* $0 < v < 1$ *and* $t > 0$*, we have*

$$rA\sharp_v B + (1 - r)A\nabla_v B \ge A!_v B.$$

3.2 Heron means for positive operators

The arithmetic–geometric mean inequality has been discussed in various extensions. One of them is the Heron mean, which interpolates between the arithmetic mean and the geometric mean. That is, for a fixed $v \in [0,1]$, the **Heron mean** for $A, B > 0$ and $r \in \mathbb{R}$ is defined by

$$H_r^v(A,B) = rA\sharp_v B + (1-r)(A\nabla_v B).$$

As a weighted version of (ii) in Proposition 3.1.1, we can easily show that

$$rA\sharp_v B + (1-r)A\nabla_v B \geq A!_v B \quad (r \leq 1, v \in [0,1]),$$

because we have

$$rA\sharp_v B + (1-r)A\nabla_v B \geq rA\sharp_v B + (1-r)A\sharp_v B = A\sharp_v B \geq A!_v B.$$

On the other hand, a similar generalization of (i) in Proposition 3.1.1 does not hold in general. If we take $r = 2$, $v = \frac{2}{3}$, $A = 1$ and $B = 2$, then the following scalar inequality does not hold in general:

$$2t^v \leq 1 - v + vt + \left(1 - v + vt^{-1}\right)^{-1}. \tag{3.2.1}$$

Thus our interest is to find a constant r_v such that $H_r^v(A,B) \leq A!_v B$ for $r \geq r_v$. We remark that the inequality (3.2.1) holds for $0 \leq v \leq 1/2$ and $t \geq 1$, or $1/2 \leq v \leq 1$ and $0 < t \leq 1$, [66, Lemma 2.3]. We first discuss on optimality of $r \in \mathbb{R}$ in the inequalities in Proposition 3.1.1.

Proposition 3.2.1. *The conditions on r in Proposition* 3.1.1 *are optimal. That is, we have*

$$\inf\{r \mid rA\sharp B + (1-r)A\nabla B \leq A!B\} = 2, \quad \sup\{r \mid rA\sharp B + (1-r)A\nabla B \geq A!B\} = 1.$$

Proof. Note that $rA\sharp B + (1-r)A\nabla B \leq$ (resp. $\geq$)$A!B$ is equivalent to $r \geq$ (resp., $\leq$)$\frac{(1+\sqrt{a})^2}{1+a}$. By $1 \leq \frac{(1+\sqrt{a})^2}{1+a} \leq 2$, $\lim\limits_{a\to 0} \frac{(1+\sqrt{a})^2}{1+a} = 1$ and $\lim\limits_{a\to 1} \frac{(1+\sqrt{a})^2}{1+a} = 2$, we have the conclusion. □

We study this from another viewpoint. We set

$$R(t) = \frac{t + 1 - \frac{4t}{t+1}}{t + 1 - 2\sqrt{t}}.$$

Then we have $R(0) = 1 \leq R(t) \leq 2 = R(1)$ for $t \geq 0$ and $R'(t) = \frac{1-t}{\sqrt{t}(t+1)^2}$. Since $R(t) = 1 + \frac{2\sqrt{t}}{t+1} = \frac{(1+\sqrt{t})^2}{1+t}$, we have $r \geq \frac{(1+\sqrt{a})^2}{1+a} \Leftrightarrow r \geq R(a)$.

If we put $C = A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$, then $R(C) \leq r$ if and only if $H_r(A,B) \leq A!B$. We have an another variant of Proposition 3.1.1.

Proposition 3.2.2. *If either (i) $0 \le v \le \frac{1}{2}$ and $0 < A \le B$ or (ii) $\frac{1}{2} \le v \le 1$ and $0 < B \le A$, then*

$$2A\sharp_v B \le A\nabla_v B + B!_v A.$$

If either (i′) $0 \le v \le \frac{1}{2}$ and $0 < B \le A$ or (ii′) $\frac{1}{2} \le v \le 1$ and $0 < A \le B$, then

$$2A\sharp_{1-v} B \le A\nabla_v B + B!_v A, \quad A\sharp B \le A\sharp_v B.$$

The proof is reduced to the following lemma by Kubo–Ando theory.

Lemma 3.2.1. *Put $f(t) = (1-v) + vt + ((1-v)t^{-1} + v)^{-1} - 2t^v$. Then*

$$f(t) = (1-v) + vt + \frac{t}{(1-v) + vt} - 2t^v \ge 2\sqrt{t} - 2t^v = 2\sqrt{t}(1 - t^{v-\frac{1}{2}}).$$

If either (i) $0 \le v \le \frac{1}{2}$ and $t \ge 1$ or (ii) $\frac{1}{2} \le v \le 1$ and $t \le 1$, then

$$2t^v \le (1-v) + vt + \left((1-v)t^{-1} + v\right)^{-1}.$$

If either (i′) $0 < v \le \frac{1}{2}$ and $0 < t \le 1$ or (ii′) $\frac{1}{2} \le v \le 1$ and $t \ge 1$, then

$$2t^{1-v} \le (1-v) + vt + \left((1-v)t^{-1} + v\right)^{-1}, \quad \sqrt{t} \le t^v.$$

We omit the proofs of Proposition 3.2.2 and Lemma 3.2.1. See [55] for them.

Remark 3.2.1. We find that the sign of $f(t)$ is not definite for the following cases:
(a) $v \in [0, \frac{1}{2})$ and $0 \le t \le 1$;
(b) $v \in (\frac{1}{2}, 1]$ and $t \ge 1$.

See [55] for the details.

Next, we consider a generalization of $R(t)$ by defining $R_v(t)$ for $v \in (0,1)$;

$$R_v(t) = \frac{1 - v + vt - \frac{t}{(1-v)t+v}}{1 - v + vt - t^v} \quad (t \ge 0).$$

It is clear that $R_{1/2}(t) = R(t)$, and that $R_v(0) = 1$, $R_v(1) = 2$ for all $v \in (0,1)$ and $\max\{R(t) : t \ge 0\} = 2$.

Problem 3.2.1. For $t \ge 0$, find the maximum value of $R_v(t)$.

For this problem, we pose an *answer*, as an upper bound of $R_v(t)$ for $t \ge 0$. And we note the following lemma mentioned in [66, Lemma 2.3].

Lemma 3.2.2. *If either (i) $0 \le v \le \frac{1}{2}$ and $t \ge 1$, or (ii) $\frac{1}{2} \le v \le 1$ and $0 < t \le 1$, then $2t^v \le 1 - v + vt + (1 - v + \frac{v}{t})^{-1}$.*

To solve Problem 3.2.1, we define the function $f_r(t)$ for $r \geq 2$ and $v \in [1/2, 1]$ by

$$\begin{aligned} f_r(t) &= (r-1)(1-v+vt) + \left(1-v+vt^{-1}\right)^{-1} - rt^v \\ &= (r-1)(1-v+vt) + \frac{t}{(1-v)t+v} - rt^v. \end{aligned}$$

It is easily seen that for a fixed $t > 0$, $f_r(t)$ is increasing for $r \geq 2$. As a consequence, it follows that if $v \in [1/2, 1]$, then

$$f_r(t) \geq 0 \quad \text{for } r \geq 2 \text{ and } 0 < t \leq 1. \tag{3.2.2}$$

We have the following generalized result.

Theorem 3.2.1 ([55]). *For a fixed $v \in [1/2, 1)$, if $r \geq r_v := \frac{2(2-v)}{3(1-v)}$, then we have*

$$rA\sharp_v B + (1-r)A\nabla_v B \leq A!_v B$$

for all $A, B > 0$.

See [55] for the proof of Theorem 3.2.1.

Remark 3.2.2. Related to the assumption $r \geq r_v$ in Theorem 3.2.1, we have $r_v \geq r_{1/2} = 2$, since $r_v = \frac{2}{3}(1 + \frac{1}{1-v})$ is increasing as a function of $v \in [1/2, 1)$.

Next, we consider the above argument under the strict operator order. For convenience, we denote $m_X = \min Sp(X) = \|X^{-1}\|^{-1}$ for $X > 0$.

Theorem 3.2.2 ([55]). *For given $A, B > 0$, put $C = A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$.*

(i) *If $0 \leq v \leq \frac{1}{2}$ and $B - A \geq m > 0$, then*

$$c_1 \leq A\nabla_v B + B!_v A - 2A\sharp_v B,$$

where

$$c_1 = 2m_A\left\{\left(1 + m\|A\|^{-1}\right)^{\frac{1}{2}} - \left(1 + m\|A\|^{-1}\right)^v\right\} \geq 0.$$

(ii) *If $\frac{1}{2} \leq v \leq 1$ and $A - B \geq m > 0$, then*

$$\min\{c_2, c_3\} \leq A\nabla_v B + B!_v A - 2A\sharp_v B,$$

where

$$c_2 = 2m_A\left\{\left(1 - m\|A\|^{-1}\right)^{\frac{1}{2}} - \left(1 - m\|A\|^{-1}\right)^v\right\} \geq 0, \quad c_3 = 2m_A\left\{(m_C)^{\frac{1}{2}} - (m_C)^v\right\} \geq 0.$$

To prove it, we prepare the following lemma. The proofs are not difficult computations so that we omit them. See [55].

Lemma 3.2.3. *For a fixed $v \in [0, 1]$, define $g_v(t) = 2(\sqrt{t} - t^v)$ for $t \geq 0$. Then*

(i) *If* $v \in [0, \frac{1}{2}]$, *then* g_v *is increasing on* $[1, \infty)$.
(ii) *If* $v \in [\frac{1}{2}, 1]$, *then* g_v *is decreasing on* $[1, \infty)$.
(iii) *If* $v \in [\frac{1}{2}, 1]$, *then* g_v *is concave on* $[0, 1]$.
(iv) *If* $v \in [0, \frac{1}{2}]$, *then* g_v *is increasing on* $[0, t_0]$ *and decreasing on* $[t_0, 1]$ *for some* $t_0 \in (0, 1)$.
(v) $g_v(t) \geq 0$ *for the following conditions (c) or (d).*
  (c) $v \in [0, \frac{1}{2}]$ *and* $t \geq 1$
  (d) $v \in [\frac{1}{2}, 1]$ *and* $0 \leq t \leq 1$.

Proof of Theorem 3.2.2.
(i) Since $B \geq A + m$, we have

$$C = A^{-\frac{1}{2}} B A^{-\frac{1}{2}} \geq A^{-\frac{1}{2}} (A + m) A^{-\frac{1}{2}} = 1 + mA^{-1} \geq 1 + m\|A\|^{-1}.$$

So it follows from (i) of Lemma 3.2.3 that if $t \in Sp(C)$, then g_v is increasing for $t \geq 1$ by $0 \leq v \leq \frac{1}{2}$. Hence we have $g_v(t) \geq g_v(1 + m\|A\|^{-1})$, so that by Lemma 3.2.1,

$$(1 - v) + vt + \frac{t}{(1 - v) + vt} - 2t^v \geq g_v(t) \geq g_v(1 + m\|A\|^{-1}).$$

Thus we have

$$(1 - v) + vC + \frac{C}{(1 - v) + vC} - 2C^v \geq g_v(1 + m\|A\|^{-1}),$$

which implies $c_1 \leq A \nabla_v B + B!_v A - 2A \sharp_v B$.
(ii) Since $A - m \geq B$, we have

$$m_C \leq C \leq A^{-\frac{1}{2}} (A - m) A^{-\frac{1}{2}} = 1 - mA^{-1} \leq 1 - m\|A\|^{-1}.$$

Now g_v is concave on $[0, 1]$ by (iii) of Lemma 3.2.3, so that if $t \in Sp(C)$, then

$$\min\{g_v(m_C), g_v(1 - m\|A\|^{-1})\} \leq g_v(t).$$

Therefore, we have the following inequality by the similar way to the proof of (i),

$$\min\{g_v(m_C), g_v(1 - m\|A\|^{-1})\} m_A \leq A \nabla_v B + B!_v A - 2A \sharp_v B. \qquad \square$$

Remark 3.2.3. Since $C \leq \|C\|$ and g_v is decreasing on $[1, \infty)$ by (ii) of Lemma 3.2.3, we have $g_v(t) \geq g_v(\|C\|)$. This implies the following result. If $1/2 \leq v \leq 1$ and $B - A \geq 0$ (these conditions correspond to (b) in Remark 3.2.1), then

$$c_4 \leq A \nabla_v B + B!_v A - 2A \sharp_v B,$$

where

$$c_4 = 2m_A \left(\|C\|^{\frac{1}{2}} - \|C\|^v\right) \leq 0.$$

To study the bounds for $A\nabla_v B + A!_v B - 2A\sharp_v B$ instead of $A\nabla_v B + B!_v A - 2A\sharp_v B$, we give the following lemma.

Lemma 3.2.4. *For a fixed $v \in [0,1]$, we define*

$$f_v(t) = (1-v) + vt + \frac{t}{(1-v)t+v} - 2t^v$$

for $t \geq 0$. $g_v(t)$ is defined for $t \geq 0$ in Lemma 3.2.3. We set $t_v = \frac{v(2-v)}{(1-v)(1+v)}$. Then we have the following properties:

(i) $f_v(1) = 0$. *In addition, $f_v(t) \geq 0$ for the following conditions (c) or (d).*
  (c) $v \in [0, \frac{1}{2}]$ *and* $t \geq 1$
  (d) $v \in [\frac{1}{2}, 1]$ *and* $0 \leq t \leq 1$.

(ii) *If $v \in [0, \frac{1}{2}]$, then f_v is increasing on $[1, \infty)$.*

(iii) *If $v \in [\frac{1}{2}, 1]$, then f_v is decreasing on $[0, 1]$.*

(iv) *If $v \in [0, \frac{1}{2}]$, then f_v is increasing on $[t_v, 1)$.*

(v) *If $v \in [\frac{1}{2}, 1)$, then f_v is decreasing on $(1, t_v]$.*

(vi) *If $v \in [0, \frac{1}{2}]$, then f_v is convex on $[1, \infty)$.*

(vii) *If $v \in [\frac{1}{2}, 1]$, then f_v is convex on $[0, 1]$.*

(viii) *If $v \in [0, \frac{1}{2}]$, then f_v is concave on $[t_v, 1)$.*

(ix) *If $v \in [\frac{1}{2}, 1)$, then f_v is concave on $(1, t_v]$.*

(x) *If $v \in [0, \frac{1}{2}]$ and $0 \leq t \leq 1$, then $g_v(t) \leq \min\{f_v(t), 0\}$.*

(xi) *If $v \in [\frac{1}{2}, 1]$ and $t \geq 1$, then $g_v(t) \leq \min\{f_v(t), 0\}$.*

The proofs are not difficult computations so that we omit them. See [55] for the proof of Lemma 3.2.4.

Remark 3.2.4. We find that the sign of $f_v(t)$ is not definite for the following cases:

(i) $v \in [0, 1/2)$ and $0 \leq t \leq 1$,

(ii) $v \in (1/2, 1]$ and $t \geq 1$.

Theorem 3.2.3 ([55]). *For $A, B > 0$, put $C = A^{-\frac{1}{2}} B A^{-\frac{1}{2}}$. $f_v(t)$ is defined in Lemma 3.2.4. Then we have the following inequalities:*

(i) *If $v \in [0, 1/2]$ and $B - A \geq m > 0$, then*

$$0 \leq f_v(1 + m\|A\|^{-1}) m_A \leq A\nabla_v B + A!_v B - 2A\sharp_v B \leq f_v(\|C\|)\|A\|.$$

(ii) *If $v \in [1/2, 1]$ and $A - B \geq m > 0$, then*

$$0 \leq f_v(1 - m\|A\|^{-1}) m_A \leq A\nabla_v B + A!_v B - 2A\sharp_v B \leq f_v(m_C)\|A\|.$$

Proof.

(i) Since $B \geq A + m$, we have

$$\|C\| \geq C = A^{-\frac{1}{2}} B A^{-\frac{1}{2}} \geq A^{-\frac{1}{2}}(A+m)A^{-\frac{1}{2}} \geq 1 + mA^{-1} \geq 1 + m\|A\|^{-1}.$$

From (ii) of Lemma 3.2.4, $f_v(t)$ is increasing for $t \geq 1$ when $v \in [0, 1/2]$ and $t \geq 1$ for $t \in Sp(C)$, we have $f_v(1 + m\|A\|^{-1}) \leq f_v(t) \leq f_v(\|C\|)$. Thus we have

$$f_v(1 + m\|A\|^{-1})m_A \leq A\nabla_v B + A!_v B - 2A\sharp_v B \leq f_v(\|C\|)\|A\|.$$

(ii) Since $A - m \geq B$, we have $m_C \leq C \leq 1 - m\|A\|^{-1}$. From (iii) of Lemma 3.2.4, $f_v(t)$ is decreasing for $0 < t \leq 1$ when $v \in [1/2, 1]$ and $0 < t \leq 1$ for $t \in Sp(C)$, we have $f_v(1 - m\|A\|^{-1}) \leq f_v(t) \leq f_v(m_C)$. Thus we have

$$f_v(1 - m\|A\|^{-1})m_A \leq A\nabla_v B + A!_v B - 2A\sharp_v B \leq f_v(m_C)\|A\|. \qquad \square$$

Theorem 3.2.3 give the refinements for the second inequality in [66, Theorem 2.1]. We show the following results by the similar way to the proof of Theorem 3.2.3. The conditions in (i) and (ii) of the following proposition correspond to those in (i) and (ii) of Remark 3.2.4.

Proposition 3.2.3. *For $A, B > 0$, put $C = A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$. $f_v(t)$ and t_v are defined in Lemma 3.2.4. $g_v(t)$ is also defined in Lemma 3.2.3. c_3 and c_4 are given in Theorem 3.2.2 and Remark 3.2.3. Then we have the following inequalities:*

(i) *For a given $v \in [0, 1/2]$, if $t_v A \leq B \leq A - m$ with $m > 0$, then*

$$f_v(m_C)m_A \leq A\nabla_v B + A!_v B - 2A\sharp_v B \leq f_v(1 - m\|A\|^{-1})\|A\|.$$

In particular,

$$0 \leq c_3 \leq A\nabla_v B + A!_v B - 2A\sharp_v B.$$

(ii) *For a given $v \in [1/2, 1)$, if $A + m \leq B \leq t_v A$ with $m > 0$, then*

$$f_v(||C||)m_A \leq A\nabla_v B + A!_v B - 2A\sharp_v B \leq f_v(1 + m||A||^{-1})||A||.$$

In particular,

$$c_4 \leq A\nabla_v B + A!_v B - 2A\sharp_v B.$$

Proof.

(i) The condition $t_v A \leq B \leq A - m$ implies $t_v \leq C \leq 1$ and $m_C \leq C \leq 1 - m||A||^{-1}$ as $t_v \leq m_C$. From (iv) of Lemma 3.2.4, f_v is increasing for $t_v \leq t \leq 1$ so that we have $f_v(m_C) \leq f_v(t) \leq f_v(1 - m||A||^{-1})$, since $t_v \leq t \leq 1$ if $t \in Sp(C)$. Thus we have

$$f_v(m_C)m_A \leq A\nabla_v B + A!_v B - 2A\sharp_v B \leq f_v(1 - m||A||^{-1})||A||.$$

From (x) of Lemma 3.2.4, we especially have $c_3 \leq A\nabla_v B + A!_v B - 2A\sharp_v B$.

(ii) The condition $A + m \le B \le t_v A$ implies $1 \le C \le t_v$ and $1 + m||A||^{-1} \le C \le ||C|| \le t_v$. From (v) of Lemma 3.2.4, f_v is decreasing for $1 \le t \le t_v$ so that we have $f_v(||C||) \le f_v(t) \le f_v(1 + m||A||^{-1})$, since $1 \le t \le t_v$ if $t \in Sp(C)$. Thus we have

$$f_v(||C||)m_A \le A\nabla_v B + A!_v B - 2A\sharp_v B \le f_v(1 + m||A||^{-1})||A||.$$

From (xi) of Lemma 3.2.4, we especially have $c_4 \le A\nabla_v B + A!_v B - 2A\sharp_v B$. □

Related to the strict positivity of operators, the arithmetic–geometric mean inequality is refined as follows.

Theorem 3.2.4 ([55]). *If $A - B$ is invertible for $A, B > 0$, then for each $v \in (0, 1)$,*

$$A\nabla_v B - A\sharp_v B > 0.$$

In particular, if $A - B \ge m > 0$, then

$$s_v\left(1 - \frac{m}{\|A\|}\right)m_A \le A\nabla_v B - A\sharp_v B \le s_v\left(\frac{m_B}{\|A\|}\right)\|A\|,$$

where $s_v(x) = 1 - v + vx - x^v$.

Proof. Put $C = A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$. Since $1 \notin Sp(C)$, we have

$$A\nabla_v B - A\sharp_v B = A^{\frac{1}{2}}s_v(C)A^{\frac{1}{2}} \ge \epsilon A \ge \epsilon m_A$$

for some $\epsilon > 0$.

Next, if $A - B \ge m > 0$, then C has bounds such that

$$\frac{m_B}{\|A\|} \le m_B A^{-1} \le C \le (A - m)A^{-1} = 1 - mA^{-1} \le 1 - \frac{m}{\|A\|} < 1.$$

Noting that s_v is convex, decreasing and $s_v(x) > 0$ on $[0, 1)$, we have

$$s_v\left(\frac{m_B}{\|A\|}\right) \ge s_v(C) \ge s_v\left(1 - \frac{m}{\|A\|}\right).$$

Since $A\nabla_v B - A\sharp_v B = A^{\frac{1}{2}}s_v(C)A^{\frac{1}{2}}$,

$$s_v\left(1 - \frac{m}{\|A\|}\right)m_A \le A\nabla_v B - A\sharp_v B \le s_v\left(\frac{m_B}{\|A\|}\right)\|A\|,$$

as desired. □

Lemma 3.2.5. *If $A - B \ge m$ for some $m > 0$, then $B^{-1} - A^{-1} \ge \frac{m}{(\|B\|+m)\|B\|} =: m_1$.*

It is easily proved as

$$B^{-1} - A^{-1} \ge B^{-1} - (B + m)^{-1} = mB^{-1}(B + m)^{-1} \ge m_1.$$

By the use of Lemma 3.2.5, we have a refinement of the geometric-harmonic mean inequality.

Corollary 3.2.1. *Notation as in above. If $A - B \geq m$ for some $m > 0$ and $0 < v < 1$, then*

$$A\sharp_v B - A!_v B \geq \frac{m_2}{(\|B_1\| + m_2)\|B_1\|},$$

where $B_1 = A\sharp_v B$ and $m_2 := s_{1-v}(1 - \frac{m_1}{\|B^{-1}\|})m_B$.

Proof. Combining Lemma 3.2.5 with Theorem 3.2.4, we have

$$B^{-1}\nabla_{1-v}A^{-1} - B^{-1}\sharp_{1-v}A^{-1} \geq s_{1-v}\left(1 - \frac{m_1}{\|B^{-1}\|}\right)m_B = m_2.$$

If we put $B_1 = (A\sharp_v B)^{-1} = B^{-1}\sharp_{1-v}A^{-1}$, then it follows from Lemma 3.2.5 that

$$A\sharp_v B - A!_v B = \left(B^{-1}\sharp_{1-v}A^{-1}\right)^{-1} - \left(B^{-1}\nabla_{1-v}A^{-1}\right)^{-1} \geq \frac{m_2}{(\|B_1\| + m_2)\|B_1\|},$$

as desired. □

3.3 Inequalities related to Heron mean

Applying the results given in Section 2.6, we prove new inequalities. To state our result, we recall that the family of **Heron mean** [21] for $a, b > 0$ is defined as

$$H_r^v(a,b) = ra\sharp_v b + (1-r)a\nabla_v b, \quad v \in [0,1] \quad \text{and} \quad r \in \mathbb{R}.$$

The author [66] showed that if $r \leq 1$, then

$$a!_v b \leq H_r^v(a,b), \quad v \in [0,1]. \tag{3.3.1}$$

Theorem 3.3.1 ([79]). *Let $a, b > 0$, $r \in \mathbb{R}$ and $0 \leq v \leq 1$. We define the following functions:*

$$g_{r,v}(a,b) = v\left(\frac{b-a}{a}\right)\left\{r\left(\frac{a+b}{2a}\right)^{v-1} + (1-r)\right\} + 1,$$

$$G_{r,v}(a,b) = \frac{v}{2}\left(\frac{b-a}{a}\right)\left(ra^{1-v}b^{v-1} + 2 - r\right) + 1.$$

(i) *If either $a \leq b$, $r \geq 0$ or $b \leq a$, $r \leq 0$, then $g_{r,v}(a,b) \leq H_r^v(a,b) \leq G_{r,v}(a,b)$.*
(ii) *If either $a \leq b$, $r \leq 0$ or $b \leq a$, $r \geq 0$, then $G_{r,v}(a,b) \leq H_r^v(a,b) \leq g_{r,v}(a,b)$.*

Our main idea and technical tool on the inequalities in Theorem 3.3.1 are closely related to the Hermite–Hadamard inequality (2.6.5).

Proof. Consider the function $f_{r,v}(t) = rvt^{v-1} + (1-r)v$ where $t > 0$, $r \in \mathbb{R}$ and $v \in [0,1]$. Since the function $f_{r,v}(t)$ is twice differentiable, one can easily see that

$$\frac{df_{r,v}(t)}{dt} = r(v-1)vt^{v-2}, \quad \frac{d^2 f_{r,v}(t)}{dt^2} = r(v-2)(v-1)vt^{v-3}.$$

It is not hard to check that

$$\begin{cases} \frac{d^2 f_{r,v}(t)}{dt^2} \geq 0 & \text{for } r \geq 0 \\ \frac{d^2 f_{r,v}(t)}{dt^2} \leq 0 & \text{for } r \leq 0. \end{cases}$$

Applying the Hermite–Hadamard inequality (2.6.5) for the function $f_{r,v}(t)$ we infer that

$$g_{r,v}(x) \leq rx^v + (1-r)\big((1-v)+vx\big) \leq G_{r,v}(x), \tag{3.3.2}$$

where

$$g_{r,v}(x) = v(x-1)\left\{r\left(\frac{1+x}{2}\right)^{v-1} + (1-r)\right\} + 1, \tag{3.3.3}$$

$$G_{r,v}(x) = \frac{v(x-1)}{2}\big(rx^{v-1} + 2 - r\big) + 1, \tag{3.3.4}$$

for each $x \geq 1$, $r \geq 0$ and $v \in [0,1]$. Similarly, for each $0 < x \leq 1$, $r \geq 0$ and $v \in [0,1]$, we get

$$G_{r,v}(x) \leq rx^v + (1-r)\big((1-v)+vx\big) \leq g_{r,v}(x). \tag{3.3.5}$$

If $x \geq 1$ and $r \leq 0$, we get

$$G_{r,v}(x) \leq rx^v + (1-r)\big((1-v)+vt\big) \leq g_{r,v}(x), \tag{3.3.6}$$

for each $v \in [0,1]$. For the case $0 < x \leq 1$ and $r \leq 0$, we have

$$g_{r,v}(t) \leq rx^v + (1-r)\big((1-v)+vt\big) \leq G_{r,v}(x), \tag{3.3.7}$$

for each $v \in [0,1]$. □

Note that we equivalently obtain the operator inequalities from the scalar inequalities given in Theorem 3.3.1 by the Kubo–Ando theory. We here omit such expressions for simplicity.

Closing this section, we prove the ordering

$$\{(1-v)+vt^{-1}\}^{-1} \leq g_{r,v}(t), \quad \text{and} \quad \{(1-v)+vt^{-1}\}^{-1} \leq G_{r,v}(t)$$

under some assumptions, for the purpose to show the advantages of our lower bounds given in Theorem 3.3.1. It is known that

$$\{(1-v)+vt^{-1}\}^{-1} \leq t^v, \quad v \in [0,1] \text{ and } t > 0,$$

so that we also have interests in the ordering $g_{r,v}(t)$ and $G_{r,v}(t)$ with t^v. That is, we can show the following four propositions.

Proposition 3.3.1. *For $t \ge 1$ and $0 \le v, r \le 1$, we have*

$$\{(1-v)+vt^{-1}\}^{-1} \le g_{r,v}(t). \tag{3.3.8}$$

Proposition 3.3.2. *For $0 < t \le 1$ and $0 \le v, r \le 1$, we have*

$$\{(1-v)+vt^{-1}\}^{-1} \le t^v \le g_{r,v}(t). \tag{3.3.9}$$

Proposition 3.3.3. *For $0 \le r, v \le 1$ and $c \le t \le 1$ with $c = \frac{2^7-1}{5^4}$, we have*

$$\{(1-v)+vt^{-1}\}^{-1} \le G_{r,v}(t). \tag{3.3.10}$$

Proposition 3.3.4. *For $0 \le v \le 1$, $r \le 1$ and $t \ge 1$, we have*

$$\{(1-v)+vt^{-1}\}^{-1} \le t^v \le G_{r,v}(t). \tag{3.3.11}$$

The proofs are given in the following.

Proof of Proposition 3.3.1. Since $g_{r,v}(t)$ is decreasing in r, $g_{r,v}(t) \ge g_{1,v}(t)$ so that we have only to prove for $t \ge 1$ and $0 \le v \le 1$, the inequality $g_{1,v}(t) \ge \{(1-v)+vt^{-1}\}^{-1}$ which is equivalent to the inequality by $v(t-1) \ge 0$,

$$\left(\frac{t+1}{2}\right)^{v-1} \ge \frac{1}{(1-v)t+v}. \tag{3.3.12}$$

Since $t \ge 1$ and $0 \le v \le 1$, we have $t(\frac{t+1}{2})^{v-1} \ge t^v$. In addition, for $t > 0$ and $0 \le v \le 1$, we have $t^v \ge \{(1-v)+vt^{-1}\}^{-1}$. Thus we have $t(\frac{t+1}{2})^{v-1} \ge \{(1-v)+vt^{-1}\}^{-1}$ which implies the inequality (3.3.8). □

Proof of Proposition 3.3.2. The first inequality is known for $t > 0$ and $0 \le v \le 1$. Since $g_{r,v}(t)$ is deceasing in r, in order to prove the second inequality we have only to prove $g_{1,v}(t) \ge t^v$, that is,

$$v(t-1)\left(\frac{t+1}{2}\right)^{v-1} + 1 \ge t^v,$$

which is equivalent to the inequality

$$\frac{t^v-1}{v} \le (t-1)\left(\frac{t+1}{2}\right)^{v-1}.$$

By the use of the **Hermite–Hadamard inequality** with a convex function x^{v-1} for $0 \le v \le 1$ and $x > 0$, the above inequality can be proven as

$$\left(\frac{t+1}{2}\right)^{v-1} \le \frac{1}{1-t}\int_t^1 x^{v-1}\,dx = \frac{1-t^v}{v(1-t)}.$$ □

Proof of Proposition 3.3.3. We first prove $h(t) = 2(t-1) - \log t \geq 0$ for $c \leq t \leq 1$. Since $h''(t) \geq 0$, $h(1) = 0$ and $h(c) \approx -0.0000354367 < 0$. Thus we have $h(t) \leq 0$ for $c \leq t \leq 1$. Second, we prove $l_v(t) = 2(t-1) - ((1-v)t + v)\log t \leq 0$. Since $\frac{dl_v(t)}{dv} = (t-1)\log t \geq 0$, we have $l_v(t) \leq l_1(t) = h(t) \leq 0$. Since $G_{r,v}(t)$ is decreasing in r, we have $G_{r,v}(t) \geq G_{1,v}(t)$ so that we have only to prove $G_{1,v}(t) \geq \{(1-v) + vt^{-1}\}^{-1}$, which is equivalent to the inequality, by $v(t-1) \leq 0$,

$$\frac{t^{v-1}+1}{2} \leq \frac{1}{(1-v)t+v}$$

for $0 \leq r, v \leq 1$ and $c \leq t \leq 1$. To this end, we set $f_v(t) = 2 - (t^{v-1}+1)((1-v)t+v)$. Some calculations imply $f_v(t) \geq f_1(t) = 0$. □

Proof of Proposition 3.3.4. The first inequality is known for $t > 0$ and $0 \leq v \leq 1$. Since $G_{r,v}(t)$ is deceasing in r, in order to prove the second inequality we have only to prove $G_{1,v}(t) \geq t^v$, which is equivalent to the inequality

$$\frac{1}{2}v(t-1)\left(t^{v-1}+1\right) + 1 \geq t^v.$$

To this end, we set

$$k_v(t) = v(t-1)\left(t^{v-1}+1\right) + 2 - 2t^v.$$

Some calculations imply $k_v(t) \geq k_v(1) = 0$. □

Remark 3.3.1. Propositions 3.3.1–3.3.4 show that lower bounds given in Theorem 3.3.1 are tighter than the known bound (Harmonic mean), for the cases given in Propositions 3.3.1–3.3.4. If $r = 1$ in Proposition 3.3.1, then $g_{r,v}(t) \leq t^v$, for $t \geq 1$ and $0 \leq v \leq 1$. If $r = 1$ in Proposition 3.3.3, then $G_{r,v}(t) \leq t^v$, for $c \leq t \leq 1$ and $0 \leq v \leq 1$. We thus find that Proposition 3.3.1 and Proposition 3.3.3 make sense for the purpose of finding the functions between $\{(1-v) + vt^{-1}\}^{-1}$ and t^v.

Remark 3.3.2. In the process of the proof in Proposition 3.3.3, we find the inequality:

$$\frac{t^v + t}{2} \leq \left\{(1-v) + vt^{-1}\right\}^{-1},$$

for $0 \leq v \leq 1$ and $c \leq t \leq 1$. Then we have the following inequalities:

$$\frac{A\sharp_v B + B}{2} \leq A!_v B \leq A\sharp_v B,$$

for $0 < cA \leq B \leq A$ with $c = \frac{2^7-1}{5^4}$, and $0 \leq v \leq 1$.

In the process of the proof in Proposition 3.3.2 we also find the inequality:

$$t\left(\frac{t+1}{2}\right)^{v-1} \leq \left\{(1-v) + vt^{-1}\right\}^{-1},$$

for $0 \le v \le 1$ and $0 \le t \le 1$. Then we have the following inequalities:

$$BA^{-1/2}\left(\frac{A^{-1/2}BA^{-1/2}+I}{2}\right)^{v-1}A^{1/2} \le A!_v B \le A\sharp_v B,$$

for $0 < B \le A$ and $0 \le v \le 1$.

3.4 Some refinements of operator inequalities for positive linear maps

According to the celebrated result by M. Lin [142, Theorem 2.1], if $0 < m \le A, B \le M$, then

$$\Phi^2(A\nabla B) \le K^2(h)\Phi^2(A\sharp B), \tag{3.4.1}$$

and

$$\Phi^2(A\nabla B) \le K^2(h)\big(\Phi(A)\sharp\Phi(B)\big)^2, \tag{3.4.2}$$

where $h = \frac{M}{m}$. We call them **Lin's squaring inequalities**. Related to this, J. Xue and X. Hu [236, Theorem 2] proved that if $0 < m \le A \le m' < M' \le B \le M$, then

$$\Phi^2(A\nabla B) \le \frac{K^2(h)}{K(h')}\Phi^2(A\sharp B), \tag{3.4.3}$$

and

$$\Phi^2(A\nabla B) \le \frac{K^2(h)}{K(h')}\big(\Phi(A)\sharp\Phi(B)\big)^2, \tag{3.4.4}$$

where $h = \frac{M}{m}$ and $h' = \frac{M'}{m'}$.

We show a stronger result than (3.4.3) and (3.4.4) in Theorem 3.4.1 below. Lin's results was further generalized by several authors as we mentioned in Remark 2.8.1. Among them, X. Fu and C. He [47, Theorem 4] generalized (3.4.1) and (3.4.2) to the power of p ($2 \le p < \infty$) as follows:

$$\Phi^p(A\nabla B) \le \left(\frac{(M+m)^2}{4^{\frac{2}{p}}Mm}\right)^p \Phi^p(A\sharp B), \tag{3.4.5}$$

and

$$\Phi^p(A\nabla B) \le \left(\frac{(M+m)^2}{4^{\frac{2}{p}}Mm}\right)^p \big(\Phi(A)\sharp\Phi(B)\big)^p. \tag{3.4.6}$$

It is interesting to ask whether the inequalities (3.4.5) and (3.4.6) can be improved. This is an another motivation of the present section. We show Theorem 3.4.2 below. We also show an improvement of the **operator Pólya–Szegö inequality** in Theorem 3.4.5 below.

Lemma 3.4.1. *Let $A, B > 0$ and $\alpha > 0$, then*

$$A \le \alpha B \Leftrightarrow \|A^{\frac{1}{2}} B^{-\frac{1}{2}}\| \le \alpha^{\frac{1}{2}}.$$

The following is the famous **Ando inequality** [10].

Lemma 3.4.2. *Let Φ be any (not necessary normalized) positive linear map. Then for every $v \in [0,1]$ and $A, B > 0$,*

$$\Phi(A \sharp_v B) \le \Phi(A) \sharp_v \Phi(B).$$

The following basic lemma is essentially known as in [248, Theorem 7], but our expression is a little bit different from those in [248]. For the sake of convenience, we give it a slim proof.

Lemma 3.4.3. *Let $0 < m \le B \le m' < M' \le A \le M$, then*

$$K^r(h')(A^{-1} \sharp_v B^{-1}) \le A^{-1} \nabla_v B^{-1}, \tag{3.4.7}$$

for each $v \in [0,1]$ and $r = \min\{v, 1-v\}$ and $h' = \frac{M'}{m'}$.

Proof. From [248, Corollary 3], for any $a, b > 0$ and $v \in [0,1]$ we have

$$K^r(h) a^{1-v} b^v \le a \nabla_v b, \tag{3.4.8}$$

where $h = \frac{b}{a}$ and $r = \min\{v, 1-v\}$. Taking $a = 1$ and $b = x$, then utilizing the continuous functional calculus and the fact that $0 < h'I \le X \le hI$, we have

$$\min_{h' \le x \le h} K^r(x) X^v \le (1-v)I + vX.$$

Replacing X with $1 < h' = \frac{M'}{m'} \le A^{\frac{1}{2}} B^{-1} A^{\frac{1}{2}} \le \frac{M}{m} = h$

$$\min_{h' \le x \le h} K^r(x) (A^{\frac{1}{2}} B^{-1} A^{\frac{1}{2}})^v \le (1-v)I + vA^{\frac{1}{2}} B^{-1} A^{\frac{1}{2}}.$$

Since $K(x)$ is an increasing function for $x > 1$, then

$$K^r(h') (A^{\frac{1}{2}} B^{-1} A^{\frac{1}{2}})^v \le (1-v)I + vA^{\frac{1}{2}} B^{-1} A^{\frac{1}{2}}.$$

Now by multiplying both sides by $A^{-\frac{1}{2}}$, we deduce the desired inequality (3.4.7). □

Inequalities (3.4.3) and (3.4.4) can be generalized by means of the weighted parameter $v \in [0,1]$ as follows.

Theorem 3.4.1 ([182, Theorem 2.1]). *If* $0 < m \le B \le m' < M' \le A \le M$, *then for each* $v \in [0,1]$ *we have*

$$\Phi^2(A\nabla_v B) \le \left(\frac{K(h)}{K^r(h')}\right)^2 \Phi^2(A\sharp_v B), \tag{3.4.9}$$

and

$$\Phi^2(A\nabla_v B) \le \left(\frac{K(h)}{K^r(h')}\right)^2 (\Phi(A)\sharp_v \Phi(B))^2, \tag{3.4.10}$$

where $r = \min\{v, 1-v\}$, $h = \frac{M}{m}$, $h' = \frac{M'}{m'}$.

Proof. According to the hypothesis, we have $A+MmA^{-1} \le M+m$ and $B+MmB^{-1} \le M+m$. Also the following inequalities hold:

$$(1-v)A + (1-v)MmA^{-1} \le (1-v)M + (1-v)m, \tag{3.4.11}$$

and

$$vB + vMmB^{-1} \le vM + vm. \tag{3.4.12}$$

Now summing up (3.4.11) and (3.4.12) we obtain $A\nabla_v B + Mm(A^{-1}\nabla_v B^{-1}) \le M + m$. Applying a positive linear map Φ we can write

$$\Phi(A\nabla_v B) + Mm\Phi(A^{-1}\nabla_v B^{-1}) \le M + m. \tag{3.4.13}$$

With the inequality (3.4.13) in hand, we are ready to prove (3.4.9). By Lemma 3.4.1, it is enough to prove that $\|\Phi(A\nabla_v B)\Phi^{-1}(A\sharp_v B)\| \le \frac{K(h)}{K^r(h')}$. By computations, we have

$$\begin{aligned}
&\|\Phi(A\nabla_v B)MmK^r(h')\Phi^{-1}(A\sharp_v B)\| \\
&\le \frac{1}{4}\|\Phi(A\nabla_v B) + MmK^r(h')\Phi^{-1}(A\sharp_v B)\|^2 \quad \text{(by Lemma 2.6.2)} \\
&\le \frac{1}{4}\|\Phi(A\nabla_v B) + MmK^r(h')\Phi(A^{-1}\sharp_v B^{-1})\|^2 \quad \text{(by Lemma 2.6.3)} \\
&\le \frac{1}{4}\|\Phi(A\nabla_v B) + Mm\Phi(A^{-1}\nabla_v B^{-1})\|^2 \quad \text{(by Lemma 3.4.3)} \\
&\le \frac{1}{4}(M+m)^2 \quad \text{(by (3.4.13))}
\end{aligned} \tag{3.4.14}$$

which leads to (3.4.9). Now we prove (3.4.10). The operator inequality (3.4.10) is equivalent to

$$\|\Phi(A\nabla_v B)(\Phi(A)\sharp_v \Phi(B))^{-1}\| \le \frac{K(h)}{K^r(h')}.$$

Compute

$$
\begin{aligned}
&\left\|\Phi(A\nabla_v B)MmK^r(h')(\Phi(A)\sharp_v\Phi(B))^{-1}\right\| \\
&\le \frac{1}{4}\left\|\Phi(A\nabla_v B)+MmK^r(h')(\Phi(A)\sharp_v\Phi(B))^{-1}\right\|^2 \quad \text{(by Lemma 2.6.2)} \\
&\le \frac{1}{4}\left\|\Phi(A\nabla_v B)+MmK^r(h')\Phi^{-1}(A\sharp_v B)\right\|^2 \quad \text{(by Lemma 3.4.2)} \\
&\le \frac{1}{4}(M+m)^2 \quad \text{(by (3.4.14))}.
\end{aligned}
$$

We thus complete the proof. □

Remark 3.4.1. Inequalities (3.4.3) and (3.4.4) are two special cases of Theorem 3.4.1 by taking $v = 1/2$. In addition, our inequalities in Theorem 3.4.1 are tighter than that in (3.4.1) and (3.4.2).

To achieve the second result, we state for easy reference the following fact obtaining from [10, Theorem 3] that will be applied in the below.

Lemma 3.4.4 ([10, Theorem 3]). *For $A, B \ge 0$ and $1 \le r < \infty$, we have*

$$\|A^r + B^r\| \le \|(A+B)^r\|.$$

Our promised refinement of inequalities (3.4.5) and (3.4.6) can be stated as follows.

Theorem 3.4.2 ([182, Theorem 2.2]). *If $0 < m \le B \le m' < M' \le A \le M$, then for each $2 \le p < \infty$ and $v \in [0,1]$ we have*

$$\Phi^p(A\nabla_v B) \le \left(\frac{K(h)}{4^{\frac{2}{p}-1}K^r(h')}\right)^p \Phi^p(A\sharp_v B), \tag{3.4.15}$$

and

$$\Phi^p(A\nabla_v B) \le \left(\frac{K(h)}{4^{\frac{2}{p}-1}K^r(h')}\right)^p (\Phi(A)\sharp_v\Phi(B))^p, \tag{3.4.16}$$

where $r = \min\{v, 1-v\}$, $h = \frac{M}{m}$, $h' = \frac{M'}{m'}$.

Proof. It can be easily seen that the operator inequality (3.4.15) is equivalent to

$$\left\|\Phi^{\frac{p}{2}}(A\nabla_v B)\Phi^{-\frac{p}{2}}(A\sharp_v B)\right\| \le \frac{(M+m)^p}{4M^{\frac{p}{2}}m^{\frac{p}{2}}K^{\frac{pr}{2}}(h')}.$$

By simple computation,

$$\begin{aligned}
&\|\Phi^{\frac{p}{2}}(A\nabla_v B)M^{\frac{p}{2}}m^{\frac{p}{2}}K^{\frac{pr}{2}}(h')\Phi^{-\frac{p}{2}}(A\sharp_v B)\| \\
&\quad\le \frac{1}{4}\|\Phi^{\frac{p}{2}}(A\nabla_v B)+M^{\frac{p}{2}}m^{\frac{p}{2}}K^{\frac{pr}{2}}(h')\Phi^{-\frac{p}{2}}(A\sharp_v B)\|^2 \quad \text{(by Lemma 2.6.2)}\\
&\quad\le \frac{1}{4}\|(\Phi(A\nabla_v B)+MmK^r(h')\Phi^{-1}(A\sharp_v B))^{\frac{p}{2}}\|^2 \quad \text{(by Lemma 3.4.4)}\\
&\quad= \frac{1}{4}\|\Phi(A\nabla_v B)+MmK^r(h')\Phi^{-1}(A\sharp_v B)\|^p\\
&\quad\le \frac{1}{4}(M+m)^p \quad \text{(by (3.4.14))}
\end{aligned}$$

which leads to (3.4.15). The desired inequality (3.4.16) is equivalent to

$$\|\Phi^{\frac{p}{2}}(A\nabla_v B)\big(\Phi(A)\sharp_v\Phi(B)\big)^{-\frac{p}{2}}\| \le \frac{(M+m)^p}{4M^{\frac{p}{2}}m^{\frac{p}{2}}K^{\frac{pr}{2}}(h')}.$$

The result follows from

$$\begin{aligned}
&\|\Phi^{\frac{p}{2}}(A\nabla_v B)M^{\frac{p}{2}}m^{\frac{p}{2}}K^{\frac{pr}{2}}(h')(\Phi(A)\sharp_v\Phi(B))^{-\frac{p}{2}}\| \\
&\quad\le \frac{1}{4}\|\Phi^{\frac{p}{2}}(A\nabla_v B)+M^{\frac{p}{2}}m^{\frac{p}{2}}K^{\frac{pr}{2}}(h')(\Phi(A)\sharp_v\Phi(B))^{-\frac{p}{2}}\|^2 \quad \text{(by Lemma 2.6.2)}\\
&\quad\le \frac{1}{4}\|(\Phi(A\nabla_v B)+MmK^r(h')(\Phi(A)\sharp_v\Phi(B))^{-1})^{\frac{p}{2}}\|^2 \quad \text{(by Lemma 3.4.4)}\\
&\quad= \frac{1}{4}\|\Phi(A\nabla_v B)+MmK^r(h')(\Phi(A)\sharp_v\Phi(B))^{-1}\|^p\\
&\quad\le \frac{1}{4}\|\Phi(A\nabla_v B)+MmK^r(h')\Phi^{-1}(A\sharp_v B)\|^p \quad \text{(by Lemma 3.4.2)}\\
&\quad\le \frac{1}{4}(M+m)^p \quad \text{(by (3.4.14))}
\end{aligned}$$

as required. □

Remark 3.4.2. Note that the Kantorovich constant $K(h)$ is an increasing function on $h \in [1,\infty)$. Moreover $K(h) \ge 1$ for any $h > 0$. Therefore, Theorem 3.4.2 is a refinement of the inequalities, (3.4.5) and (3.4.6) for $2 \le p < \infty$.

It is proved in [10, Theorem 2.6] that for $4 \le p < \infty$,

$$\Phi^p(A\nabla B) \le \frac{(K(h)(M^2+m^2))^p}{16M^pm^p}\Phi^p(A\sharp B),$$

and

$$\Phi^p(A\nabla B) \le \frac{(K(h)(M^2+m^2))^p}{16M^pm^p}(\Phi(A)\sharp\Phi(B))^p.$$

These inequalities can be improved in the following.

Theorem 3.4.3 ([182, Theorem 2.3]). *If* $0 < m \le B \le m' < M' \le A \le M$, *then for each* $4 \le p < \infty$ *and* $v \in [0,1]$ *we have*

$$\Phi^p(A\nabla_v B) \le \left(\frac{\sqrt{K(h^2)}K(h)}{2^{\frac{4}{p}-1}K^r(h')}\right)^p \Phi^p(A\sharp_v B), \tag{3.4.17}$$

and

$$\Phi^p(A\nabla_v B) \le \left(\frac{\sqrt{K(h^2)}K(h)}{2^{\frac{4}{p}-1}K^r(h')}\right)^p (\Phi(A)\sharp_v\Phi(B))^p, \tag{3.4.18}$$

where $r = \min\{v, 1-v\}$, $h = \frac{M}{m}$ *and* $h' = \frac{M'}{m'}$.

Proof. From (3.4.9), we obtain

$$\Phi^{-2}(A\sharp_v B) \le \left(\frac{K(h)}{K^r(h')}\right)^2 \Phi^{-2}(A\nabla_v B). \tag{3.4.19}$$

From this, one can see that

$$\begin{aligned}
&\|\Phi^{\frac{p}{2}}(A\nabla_v B)M^{\frac{p}{2}}m^{\frac{p}{2}}\Phi^{-\frac{p}{2}}(A\sharp_v B)\| \\
&\le \frac{1}{4}\left\|\frac{K^{\frac{p}{4}}(h)}{K^{\frac{rp}{4}}(h')}\Phi^{\frac{p}{2}}(A\nabla_v B) + \left(\frac{K^r(h')M^2m^2}{K(h)}\right)^{\frac{p}{4}}\Phi^{-\frac{p}{2}}(A\sharp_v B)\right\|^2 \quad \text{(by Lemma 2.6.2)} \\
&\le \frac{1}{4}\left\|\left(\frac{K(h)}{K^r(h')}\Phi^2(A\nabla_v B) + \frac{K^r(h')M^2m^2}{K(h)}\Phi^{-2}(A\sharp_v B)\right)^{\frac{p}{4}}\right\|^2 \quad \text{(by Lemma 3.4.4)} \\
&= \frac{1}{4}\left\|\frac{K(h)}{K^r(h')}\Phi^2(A\nabla_v B) + \frac{K^r(h')M^2m^2}{K(h)}\Phi^{-2}(A\sharp_v B)\right\|^{\frac{p}{2}} \\
&\le \frac{1}{4}\left\|\frac{K(h)}{K^r(h')}\Phi^2(A\nabla_v B) + \frac{K(h)M^2m^2}{K^r(h')}\Phi^{-2}(A\nabla_v B)\right\|^{\frac{p}{2}} \quad \text{(by (3.4.19))} \\
&= \frac{1}{4}\left\|\frac{K(h)}{K^r(h')}(\Phi^2(A\nabla_v B) + M^2m^2\Phi^{-2}(A\nabla_v B))\right\|^{\frac{p}{2}} \\
&\le \frac{1}{4}\left(\frac{K(h)(M^2+m^2)}{K^r(h')}\right)^{\frac{p}{2}} \quad \text{(by [142, Eq.(4.7)])}
\end{aligned}$$

which implies

$$\|\Phi^{\frac{p}{2}}(A\nabla_v B)\Phi^{-\frac{p}{2}}(A\sharp_v B)\| \le \frac{1}{4}\left(\frac{K(h)(M^2+m^2)}{K^r(h')Mm}\right)^{\frac{p}{2}}.$$

This proves the inequality (3.4.17). Similarly, (3.4.18) holds by the inequality (3.4.10). □

In [187], M. S. Moslehian, R. Nakamoto and Y. Seo gave the **operator Pólya–Szegö inequality** as follows.

Theorem 3.4.4 ([187, Theorem 2.1]). *Let Φ be a positive linear map. If $0 < m_1^2 \le A \le M_1^2$ and $0 < m_2^2 \le B \le M_2^2$ for some positive real numbers $m_1 \le M_1$ and $m_2 \le M_2$, then*

$$\Phi(A)\sharp\Phi(B) \le \frac{M+m}{2\sqrt{Mm}}\Phi(A\sharp B), \tag{3.4.20}$$

where $m = \frac{m_2}{M_1}$ and $M = \frac{M_2}{m_1}$.

It is worth noting that the inequality (3.4.20) was first proved in [138, Theorem 4] for matrices (see also [29, Theorem 3]). X. Zhao et al. obtained the following result in [242] by using the same strategies of [248].

Lemma 3.4.5 ([242, Theorem 3.2]). *Let $A, B \ge 0$. If either $1 < h \le A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le h'$ or $0 < h' \le A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le h < 1$, then*

$$K^r(h)A\sharp_v B \le A\nabla_v B,$$

for all $v \in [0,1]$, where $r = \min\{v, 1-v\}$.

Now, we show a new refinement of Theorem 3.4.4 thanks to Lemma 3.4.5.

Theorem 3.4.5 ([182, Theorem 2.5]). *Let A, B be two positive operators such that $m_1^2 \le A \le M_1^2$, $m_2^2 \le B \le M_2^2$, $m = \frac{m_2}{M_1}$ and $M = \frac{M_2}{m_1}$. If $m > 1$, then*

$$\Phi(A)\sharp\Phi(B) \le \gamma\Phi(A\sharp B), \tag{3.4.21}$$

where $\gamma = \frac{M+m}{2\sqrt{MmK(h)}}$, $h = \frac{m_2^2}{M_1^2}$. Moreover, the inequality (3.4.21) holds for $M < 1$ and $h = \frac{M_2^2}{m_1^2}$.

Proof. According to the assumptions, we have $m^2 = \frac{m_2^2}{M_1^2} \le A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le \frac{M_2^2}{m_1^2} = M^2$. We thus have

$$m \le (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}} \le M. \tag{3.4.22}$$

Therefore, (3.4.22) implies that $((A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}} - m)(M - (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}}) \ge 0$. Simplifying it, we find that $(M+m)(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}} \ge Mm + A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$, multiplying both sides by $A^{\frac{1}{2}}$ to get

$$(M+m)A\sharp B \ge MmA + B. \tag{3.4.23}$$

By applying a positive linear map Φ in (3.4.23), we infer

$$(M+m)\Phi(A\sharp B) \ge Mm\Phi(A) + \Phi(B). \tag{3.4.24}$$

Utilizing Lemma 3.4.5 for the case $v = \frac{1}{2}$, $A = Mm\Phi(A)$ and $B = \Phi(B)$, and by taking into account that $m_1^2 \frac{M_2 m_2}{M_1 m_1} \leq \frac{M_2 m_2}{M_1 m_1}\Phi(A) \leq M_1^2 \frac{M_2 m_2}{M_1 m_1}$ implies $m_1^2 \leq \Phi(A) \leq M_1^2$, and $1 < \frac{m_2^2}{M_1^2} \leq \Phi^{-\frac{1}{2}}(A)\Phi(B)\Phi^{-\frac{1}{2}}(A) \leq \frac{M_2^2}{m_1^2}$ we infer

$$2\sqrt{MmK(h)}\Phi(A)\sharp\Phi(B) \leq Mm\Phi(A) + \Phi(B), \tag{3.4.25}$$

where $h = \frac{m_2^2}{M_1^2}$. Combining (3.4.24) and (3.4.25), we deduce desired result (3.4.21). The case $M < 1$ is similar; we omit the details. This completes the proof. □

The inequality (3.4.21) can be squared in the following.

Theorem 3.4.6 ([182, Theorem 2.6]). *Suppose that all the assumptions of Theorem* 3.4.5 *are satisfied. Then*

$$\left(\Phi(A)\sharp\Phi(B)\right)^2 \leq \psi\Phi^2(A\sharp B), \tag{3.4.26}$$

where

$$\psi = \begin{cases} \frac{\gamma^2(\alpha+\beta)^2}{4\alpha\beta} & \text{if } \alpha \leq t_0 \\ \frac{\gamma(\alpha+\beta)-\beta}{\alpha} & \text{if } \alpha \geq t_0, \end{cases}$$

$\alpha = m_1 m_2$, $\beta = M_1 M_2$ *and* $t_0 := \frac{2\alpha\beta}{\gamma(\alpha+\beta)}$.

Proof. According to the assumption, one can see that

$$\alpha \leq \Phi(A\sharp B) \leq \beta, \tag{3.4.27}$$

and

$$\alpha \leq \Phi(A)\sharp\Phi(B) \leq \beta, \tag{3.4.28}$$

where $\alpha = m_1 m_2$, $\beta = M_1 M_2$. The inequality (3.4.27) implies $\Phi^2(A\sharp B) \leq (\alpha+\beta)\Phi(A\sharp B) - \alpha\beta$ and (3.4.28) gives us

$$\left(\Phi(A)\sharp\Phi(B)\right)^2 \leq (\alpha+\beta)\left(\Phi(A)\sharp\Phi(B)\right) - \alpha\beta. \tag{3.4.29}$$

Thus we have

$$\begin{aligned} &\Phi^{-1}(A\sharp B)\left(\Phi(A)\sharp\Phi(A)\right)^2\Phi^{-1}(A\sharp B) \\ &\quad \leq \Phi^{-1}(A\sharp B)\left((\alpha+\beta)\left(\Phi(A)\sharp\Phi(A)\right) - \alpha\beta\right)\Phi^{-1}(A\sharp B) \quad \text{(by (3.4.29))} \\ &\quad \leq \left(\gamma(\alpha+\beta)\Phi(A\sharp B) - \alpha\beta\right)\Phi^{-2}(A\sharp B) \quad \text{(by (3.4.21))}. \end{aligned} \tag{3.4.30}$$

Consider the real function $f(t)$ on $(0,\infty)$ defined as $f(t) = \frac{\gamma(\alpha+\beta)t-\alpha\beta}{t^2}$. As a matter of fact, the inequality (3.4.30) implies that

$$\Phi^{-1}(A\sharp B)\left(\Phi(A)\sharp\Phi(B)\right)^2\Phi^{-1}(A\sharp B) \leq \max_{\alpha\leq t\leq\beta} f(t).$$

One can see that the function $f(t)$ is decreasing on $[\alpha, \beta]$. By an easy computation, we have $f'(t) = \frac{2\alpha\beta - \gamma(\alpha+\beta)t}{t^3}$. This function has an maximum point on $t_0 = \frac{2\alpha\beta}{\gamma(\alpha+\beta)}$ with the maximum value $f(t_0) = \frac{\gamma^2(\alpha+\beta)^2}{4\alpha\beta}$. Where we have

$$\max_{\alpha \le t \le \beta} f(t) \le \begin{cases} f(t_0) & \text{for } \alpha \le t_0 \\ f(\alpha) & \text{for } \alpha \ge t_0. \end{cases}$$

Note that $f(\alpha) = \frac{\gamma(\alpha+\beta)-\beta}{\alpha}$. It is striking that we can get the same inequality (3.4.26) under the condition $M < 1$. Hence the proof of Theorem 3.4.6 is complete. □

4 Norm inequalities and trace inequalities

To obtain unitarily invariant norm inequalities, we apply the method established by F. Hiai and H. Kosaki [111, 112, 109, 135]. See [20] for the basic properties of a **unitarily invariant norm** $\|\cdot\|_u$, namely $\|UAV\|_u = \|A\|_u$ for any bounded linear operator A and any unitary operators U, V. F. Hiai and H. Kosaki introduced a class of **symmetric homogeneous means** in the following [111, 112, 109, 135].

Definition 4.0.1. A continuous positive real function $M(x,y)$ for $x,y > 0$ is called a symmetric homogeneous mean if the function M satisfies the following properties:
(i) $M(x,y) = M(y,x)$.
(ii) $M(cx,cy) = cM(x,y)$ for $c > 0$.
(iii) $M(x,y)$ is nondecreasing in x,y.
(iv) $\min\{x,y\} \le M(x,y) \le \max\{x,y\}$.

It is known [111, Proposition 1.2] and [135, Proposition 2.5] that the scalar inequality $M(x,y) \le N(x,y)$ for each $x,y \ge 0$ (where $M(x,y)$ and $N(x,y)$ satisfy the condition in Definition 4.0.1) is equivalent to the Frobenius norm (Hilbert–Schmidt norm) inequality $\|M(H,K)X\|_2 \le \|N(H,K)X\|_2$ for each $H,K \ge 0$ and any $X \in B(\mathcal{H})$. This fact shows that the Frobenius norm inequality is easily obtained and may be considered a kind of commutative results. In fact, we first obtained the Frobenius norm inequality in the below, see Theorem 4.1.1, as new inequalities on bounds of the logarithmic mean.

We give powerful theorem to obtain unitarily invariant norm inequalities. In the references [111, 112, 109, 135], other equivalent conditions were given. However, we give here minimum conditions to obtain our results in the next section.

Theorem 4.0.1 ([111, 112, 109, 135]). *For two symmetric homogeneous means M and N, the following conditions are equivalent:*
(i) $\|M(S,T)X\|_u \le \|N(S,T)X\|_u$ *for any* $S,T,X \in B(\mathcal{H})$ *with* $S,T \ge 0$ *and for any unitarily invariant norm* $\|\cdot\|_u$.
(ii) *The function* $M(e^t,1)/N(e^t,1)$ *is* **positive definite function** *on* $\mathbb{R}$ *(then we denote* $M \preceq N$*), where the positive definiteness of a real continuous function* ϕ *on* $\mathbb{R}$ *means that* $[\phi(t_i - t_j)]_{i,j=1,\dots,n}$ *is positive definite for any* $t_1,\dots,t_n \in \mathbb{R}$ *with any* $n \in \mathbb{N}$.

Thanks to Theorem 4.0.1, our task to obtain unitarily invariant norm inequalities in the next section is to show the relation $M \preceq N$ which is stronger than the usual scalar inequalities $M \le N$. That is, $M(s,t) \preceq N(s,t)$ implies $M(s,t) \le N(s,t)$.

Trace norm $\mathrm{Tr}[|X|]$ is one of unitarily invariant norms, but the trace $\mathrm{Tr}[X]$ is not so. Therefore, it is interesting to study the trace inequality itself. We give our results on trace inequalities in Section 4.2.

https://doi.org/10.1515/9783110643473-004

4.1 Unitarily invariant norm inequalities

In the paper [85], we derived the tight bounds for logarithmic mean in the case of Frobenius norm, inspired by the work of L. Zou in [245].

Theorem 4.1.1 ([85]). *For any matrices S, T, X with $S, T \geq 0$, $m_1 \geq 1$, $m_2 \geq 2$ and Frobenius norm $\|\cdot\|_2$, the following inequalities hold:*

$$\begin{aligned}
\frac{1}{m_1}\left\|\sum_{k=1}^{m_1} S^{k/(m_1+1)} X T^{(m_1+1-k)/(m_1+1)}\right\|_2 &\leq \frac{1}{m_1}\left\|\sum_{k=1}^{m_1} S^{(2k-1)/2m_1} X T^{(2m_1-(2k-1))/2m_1}\right\|_2 \\
&\leq \left\|\int_0^1 S^{\nu} X T^{1-\nu}\, d\nu\right\|_2 \\
&\leq \frac{1}{m_2}\left\|\sum_{k=0}^{m_2} S^{k/m_2} X T^{(m_2-k)/m_2} - \frac{1}{2}(SX + XT)\right\|_2 \\
&\leq \frac{1}{m_2}\left\|\sum_{k=0}^{m_2-1} S^{k/(m_2-1)} X T^{(m_2-1-k)/(m_2-1)}\right\|_2 .
\end{aligned}$$

Although the Frobenius norm is only a special case of unitarily invariant norms, our bounds for the logarithmic mean have improved those in the following results by F. Hiai and H. Kosaki [110, 111].

Theorem 4.1.2 ([110, 111]). *For any bounded linear operators S, T, X with $S, T \geq 0$, $m_1 \geq 1$, $m_2 \geq 2$ and any unitarily invariant norm $\|\cdot\|_u$, the following inequalities hold:*

$$\begin{aligned}
\|S^{1/2} X T^{1/2}\|_u &\leq \frac{1}{m_1}\left\|\sum_{k=1}^{m_1} S^{k/(m_1+1)} X T^{(m_1+1-k)/(m_1+1)}\right\|_u \leq \left\|\int_0^1 S^{\nu} X T^{1-\nu}\, d\nu\right\|_u \\
&\leq \frac{1}{m_2}\left\|\sum_{k=0}^{m_2-1} S^{k/(m_2-1)} X T^{(m_2-1-k)/(m_2-1)}\right\|_u \leq \frac{1}{2}\|SX + XT\|_u .
\end{aligned}$$

In this section, we give the tighter bounds for the logarithmic mean than those by F. Hiai and H. Kosaki [110, 111] for every unitarily invariant norm. That is, we give the generalized results of Theorem 4.1.1 for the unitarily invariant norm. For this purpose, we first introduce two quantities.

Definition 4.1.1. For $\alpha \in \mathbb{R}$ and $x, y > 0$, we set

$$P_\alpha(x,y) = \begin{cases} \frac{\alpha x^\alpha (x-y)}{x^\alpha - y^\alpha}, & (x \neq y) \\ x, & (x = y) \end{cases} \quad \text{and} \quad Q_\alpha(x,y) = \begin{cases} \frac{\alpha y^\alpha (x-y)}{x^\alpha - y^\alpha}, & (x \neq y) \\ x, & (x = y). \end{cases}$$

We note that we have $Q_\alpha(y,x) = P_\alpha(x,y)$ and the following bounds of logarithmic mean with the above two means (see [85, Appendix]):

$$\begin{cases} Q_{1/m}(x,y) < LM(x,y) < P_{1/m}(x,y), & (\text{if } x > y), \\ P_{1/m}(x,y) < LM(x,y) < Q_{1/m}(x,y), & (\text{if } x < y), \end{cases}$$

where the logarithmic mean $L(\cdot,\cdot)$ was defined in (2.0.2). However, we use the notation $LM(\cdot,\cdot)$ in this section to avoid the confusion with $L_\alpha(\cdot,\cdot)$ in the below. In addition, we use the notation $AM(\cdot,\cdot)$, $GM(\cdot,\cdot)$ and $HM(\cdot,\cdot)$ for nonweighted arithmetic mean, geometric mean and harmonic mean, respectively,

$$LM(x,y) = \begin{cases} \frac{x-y}{\log x - \log y}, & (x \neq y) \\ x, & (x = y). \end{cases} \tag{4.1.1}$$

We here define a few symmetric homogeneous means using $P_\alpha(x,y)$ and $Q_\alpha(x,y)$ in the following way. Actually, we considered four standard means of $P_\alpha(x,y)$ and $Q_\alpha(x,y)$ without the weight. It may be interesting to consider four general means of $P_\alpha(x,y)$ and $Q_\alpha(x,y)$ with the weight $v \in [0,1]$ as the further studies.

Definition 4.1.2.

(i) For $|\alpha| \leq 1$ and $x \neq y$, we define

$$A_\alpha(x,y) = \frac{1}{2}P_\alpha(x,y) + \frac{1}{2}Q_\alpha(x,y) = \frac{\alpha(x^\alpha + y^\alpha)(x-y)}{2(x^\alpha - y^\alpha)}.$$

(ii) For $\alpha \in \mathbb{R}$ and $x \neq y$, we define

$$L_\alpha(x,y) = \frac{P_\alpha(x,y) - Q_\alpha(x,y)}{\log P_\alpha(x,y) - \log Q_\alpha(x,y)} = LM(x,y),$$

which is independent to α,

(iii) For $|\alpha| \leq 2$ and $x \neq y$, we define

$$G_\alpha(x,y) = \sqrt{P_\alpha(x,y)Q_\alpha(x,y)} = \frac{\alpha(xy)^{\alpha/2}(x-y)}{x^\alpha - y^\alpha}.$$

(iv) For $|\alpha| \leq 1$ and $x \neq y$, we define

$$H_\alpha(x,y) = \frac{2P_\alpha(x,y)Q_\alpha(x,y)}{P_\alpha(x,y) + Q_\alpha(x,y)} = \frac{2\alpha(xy)^\alpha}{x^\alpha + y^\alpha}\frac{(x-y)}{x^\alpha - y^\alpha},$$

and we also set $A_\alpha(x,y) = L_\alpha(x,y) = G_\alpha(x,y) = H_\alpha(x,y) = x$ for $x = y$.

We have the following relations for the above means:

$$\begin{aligned}
&A_1(x,y) = AM(x,y) = \frac{1}{2}(x+y), \quad A_0(x,y) = \lim_{\alpha\to 0} A_\alpha(x,y) = LM(x,y), \\
&G_0(x,y) = \lim_{\alpha\to 0} G_\alpha(x,y) = LM(x,y), \quad G_1(x,y) = GM(x,y) = \sqrt{xy}, \\
&G_2(x,y) = HM(x,y) = \frac{2xy}{x+y}, \quad H_0(x,y) = \lim_{\alpha\to 0} H_\alpha(x,y) = LM(x,y), \\
&H_{1/2}(x,y) = GM(x,y), \quad H_1(x,y) = HM(x,y)
\end{aligned}$$

and $H_\alpha(x,y) = G_{2\alpha}(x,y)$. In addition, the above means are written as the following geometric bridges (weighted geometric means of two means) [113]:

$$A_\alpha(x,y) = [B_\alpha(x,y)]^\alpha [S_\alpha(x,y)]^{1-\alpha}, \quad L_\alpha(x,y) = [E_\alpha(x,y)]^\alpha [S_\alpha(x,y)]^{1-\alpha},$$
$$G_\alpha(x,y) = [GM(x,y)]^\alpha [S_\alpha(x,y)]^{1-\alpha}, \quad H_\alpha(x,y) = [D_\alpha(x,y)]^\alpha [S_\alpha(x,y)]^{1-\alpha},$$

where

$$S_\alpha(x,y) = \left(\frac{\alpha(x-y)}{x^\alpha - y^\alpha}\right)^{\frac{1}{1-\alpha}}, \quad B_\alpha(x,y) = \left(\frac{x^\alpha + y^\alpha}{2}\right)^{\frac{1}{\alpha}},$$

and

$$D_\alpha(x,y) = \left(\frac{2x^\alpha y^\alpha}{x^\alpha + y^\alpha}\right)^{\frac{1}{\alpha}}, \quad E_\alpha(x,y) = \left(\frac{x^\alpha - y^\alpha}{\alpha(\log x - \log y)}\right)^{\frac{1}{\alpha}}.$$

$S_\alpha(x,y)$ and $B_\alpha(x,y)$ are called **Stolarsky mean** and **binomial mean**, respectively.

In the paper [85], as tight bounds of logarithmic mean, the scalar inequalities were shown

$$G_{1/m}(x,y) \le LM(x,y), \quad (m \ge 1), \quad LM(x,y) \le A_{1/m}(x,y), \quad (m \ge 2)$$

which equivalently implied Frobenius norm inequalities (Theorem 4.1.1). See [85, Theorem 2.2 and 3.2] for details. In this section, we give unitarily invariant norm inequalities which are general results including Frobenius norm inequalities as a special case. The functions $A_\alpha(x,y), L_\alpha(x,y), G_\alpha(x,y), H_\alpha(x,y)$ defined in Definition 4.1.2 are symmetric homogeneous means. We first give monotonicity of three means $H_\alpha(x,y)$, $G_\alpha(x,y)$ and $A_\alpha(x,y)$ for the parameter $\alpha \in \mathbb{R}$. Since we have $H_{-\alpha}(x,y) = H_\alpha(x,y)$, $G_{-\alpha}(x,y) = G_\alpha(x,y)$ and $A_{-\alpha}(x,y) = A_\alpha(x,y)$, we consider the case $\alpha \ge 0$. Then we have the following proposition.

Proposition 4.1.1.
(i) *If* $0 \le \alpha < \beta \le 1$, *then* $H_\beta \preceq H_\alpha$.
(ii) *If* $0 \le \alpha < \beta \le 2$, *then* $G_\beta \preceq G_\alpha$.
(iii) *If* $0 \le \alpha < \beta \le 1$, *then* $A_\alpha \preceq A_\beta$.

Proof.
(i) We calculate

$$\frac{H_\beta(e^t,1)}{H_\alpha(e^t,1)} = \frac{2\beta e^{\beta t}(e^t-1)}{e^{2\beta t}-1} \cdot \frac{e^{2\alpha t}-1}{2\alpha e^{\alpha t}(e^t-1)} = \frac{\beta}{\alpha} \cdot \frac{e^{\beta t}(e^{2\alpha t}-1)}{e^{\alpha t}(e^{2\beta t}-1)} = \frac{\beta}{\alpha}\frac{\sinh \alpha t}{\sinh \beta t}.$$

This is a positive definite function for the case $\alpha < \beta$, so that we have $H_\beta \preceq H_\alpha$.

(ii) The similar calculation

$$\frac{G_\beta(e^{2t},1)}{G_\alpha(e^{2t},1)} = \frac{2\beta e^{\beta t}(e^{2t}-1)}{e^{2\beta t}-1}\cdot\frac{e^{2\alpha t}-1}{2\alpha e^{\alpha t}(e^{2t}-1)} = \frac{\beta}{\alpha}\cdot\frac{e^{\beta t}(e^{2\alpha t}-1)}{e^{\alpha t}(e^{2\beta t}-1)} = \frac{\beta}{\alpha}\cdot\frac{\sinh\alpha t}{\sinh\beta t}$$

implies $G_\beta \preceq G_\alpha$.

(iii) Since the case $0 = \alpha < \beta \le 1$ follows from the limit of the case $0 < \alpha < \beta \le 1$, we may assume $0 < \alpha < \beta \le 1$. Since we have

$$\frac{A_\alpha(e^{2t},1)}{A_\beta(e^{2t},1)} = \frac{\alpha}{\beta}\cdot\frac{\sinh\beta t\cosh\alpha t}{\cosh\beta t\sinh\alpha t},$$

we calculate by the formula $\sinh(x) = 2\cosh(\frac{x}{2})\sinh(\frac{x}{2})$ repeatedly

$$\begin{aligned}\frac{\sinh\beta t\cosh\alpha t}{\cosh\beta t\sinh\alpha t} - 1 &= \frac{\sinh(\beta-\alpha)t}{\cosh\beta t\sinh\alpha t} = \frac{2\cosh(\frac{\beta-\alpha}{2}t)\sinh(\frac{\beta-\alpha}{2}t)}{\cosh\beta t\sinh\alpha t}\\ &= \lim_{n\to\infty}\frac{2^n\prod_{k=1}^{n}\cos(\frac{\beta-\alpha}{2^k}t)\sinh(\frac{\beta-\alpha}{2^n}t)}{\cosh\beta t\sinh\alpha t}.\end{aligned}$$

From [26, Proposition 4], the sufficient condition that the function $\frac{\prod_{k=1}^{n}\cosh(\frac{\beta-\alpha}{2^k}t)}{\cosh\beta t}$ is positive definite, is $\sum_{k=1}^{n}\frac{\beta-\alpha}{\beta 2^k} \le 1$, that is, $(\beta-\alpha)(1-2^{-n}) \le \beta$. The sufficient condition that the function $\frac{\sinh(\frac{\beta-\alpha}{2^n}t)}{\sinh\alpha t}$ is positive definite, is $\frac{\beta-\alpha}{2^n} \le \alpha$. These conditions $(\beta-\alpha)(1-2^{-n}) \le \beta$ and $\frac{\beta-\alpha}{2^n} \le \alpha$ are satisfied with a natural number n sufficiently large. Thus we conclude $A_\alpha \preceq A_\beta$. □

It may be notable that (iii) of the above proposition can be proven by the similar argument in [111, Theorem 2.1]. Next, we give the relation among four means $H_\alpha(x,y)$, $G_\alpha(x,y)$, $L_\alpha(x,y)$, and $A_\alpha(x,y)$.

Proposition 4.1.2. *For any $S,T,X \in \mathbb{B}(\mathcal{H})$ with $S,T \ge 0$, $|\alpha| \le 1$ and any unitarily invariant norm $\|\cdot\|_u$, we have*

$$\|H_\alpha(S,T)X\|_u \le \|G_\alpha(S,T)X\|_u \le \|L_\alpha(S,T)X\|_u \le \|A_\alpha(S,T)X\|_u.$$

Proof. We first calculate

$$\frac{H_\alpha(e^t,1)}{G_\alpha(e^t,1)} = \frac{2\alpha e^{\alpha t}}{e^{\alpha t}+1}\frac{(e^t-1)}{e^{\alpha t}-1}\frac{e^{\alpha t}-1}{\alpha e^{\alpha t/2}(e^t-1)} = \frac{2e^{\alpha t/2}}{e^{\alpha t}+1} = \frac{2}{e^{\alpha t/2}+e^{-\alpha t/2}} = \frac{1}{\cosh\frac{\alpha t}{2}},$$

which is a positive definite function. Thus we have $H_\alpha \preceq G_\alpha$ so that the first inequality of this proposition holds thanks to Theorem 4.0.1. Calculations

$$\frac{G_\alpha(e^t,1)}{L_\alpha(e^t,1)} = \frac{\alpha e^{\alpha t/2}(e^t-1)}{e^{\alpha t}-1}\cdot\frac{t}{e^t-1} = \frac{\alpha t}{e^{\alpha t/2}-e^{-\alpha t/2}} = \frac{\frac{\alpha t}{2}}{\sinh\frac{\alpha t}{2}}$$

implies $G_\alpha \preceq L_\alpha$. Thus we have the second inequality of this proposition. Finally, calculations

$$\frac{L_\alpha(e^t,1)}{A_\alpha(e^t,1)} = \frac{e^t-1}{t}\cdot\frac{2(e^{\alpha t}-1)}{\alpha(e^{\alpha t}+1)(e^t-1)} = \frac{2}{\alpha t}\cdot\frac{e^{\alpha t/2}-e^{-\alpha t/2}}{e^{\alpha t/2}+e^{-\alpha t/2}} = \frac{\tanh\frac{\alpha t}{2}}{\frac{\alpha t}{2}}$$

implies $L_\alpha \preceq A_\alpha$. Thus we have the third inequality of this proposition. □

In the papers [110, 111], the unitarily invariant norm inequalities of the **power difference mean** $M_\alpha(x,y)$ was systematically studied. We give the relation our means with the power difference mean:

$$M_\alpha(x,y) = \begin{cases} \frac{\alpha-1}{\alpha}\cdot\frac{x^\alpha-y^\alpha}{x^{\alpha-1}-y^{\alpha-1}}, & (x\neq y)\\ x, & (x=y).\end{cases}$$

Theorem 4.1.3 ([67]). *For any $S,T,X\in\mathbb{B}(\mathcal{H})$ with $S,T\geq 0$, $m\in\mathbb{N}$ and any unitarily invariant norm $\|\cdot\|_u$, we have*

$$\left\|M_{\frac{m}{m+1}}(S,T)X\right\|_u \leq \left\|G_{\frac{1}{m}}(S,T)X\right\|_u \leq \left\|L(S,T)X\right\|_u \leq \left\|A_{\frac{1}{m}}(S,T)X\right\|_u \leq \left\|M_{\frac{m+1}{m}}(S,T)X\right\|_u.$$

(More concrete expressions of these inequalities will be written down in (4.1.3) below.)

Proof. The second inequality and the third inequality have already been proven in Proposition 4.1.2. To prove the first inequality, for $0<\alpha,\beta<1$ we calculate

$$\begin{aligned}\frac{M_\beta(e^{2t},1)}{G_\alpha(e^{2t},1)} &= \frac{\beta-1}{\beta}\cdot\frac{e^{2\beta t}-1}{e^{2(\beta-1)t}-1}\cdot\frac{e^{2\alpha t}-1}{\alpha e^{\alpha t}(e^{2t}-1)} = \frac{1-\beta}{\alpha\beta}\cdot\frac{\sinh\beta t}{\sinh t}\cdot\frac{\sinh\alpha t}{\sinh(1-\beta)t}\\ &= \frac{2(1-\beta)}{\alpha\beta}\cdot\frac{\sinh\beta t\cosh\frac{\alpha t}{2}}{\sinh t}\cdot\frac{\sinh\frac{\alpha t}{2}}{\sinh(1-\beta)t}.\end{aligned}$$

By [26, Proposition 5], the function $\frac{\sinh\beta t\cosh\frac{\alpha t}{2}}{\sinh t}$ is positive definite, if $\beta+\frac{\alpha}{2}\leq 1$ and $\frac{\alpha}{2}\leq\frac{1}{2}$. The function $\frac{\sinh\frac{\alpha t}{2}}{\sinh(1-\beta)t}$ is also positive definite, if $\frac{\alpha}{2}\leq 1-\beta$. The case $\alpha=\frac{1}{m}$ and $\beta=\frac{m}{m+1}$ satisfies the above conditions. Thus we have $M_{\frac{m}{m+1}}\preceq G_{\frac{1}{m}}$ which leads to the first inequality of this theorem. To prove the last inequality, for $0<\alpha<1$ and $\beta>1$, we also calculate

$$\begin{aligned}\frac{A_\alpha(e^{2t},1)}{M_\beta(e^{2t},1)} &= \frac{\alpha\beta}{2(\beta-1)}\cdot\frac{\sinh t\sinh(\beta-1)t}{\tanh\alpha t\sinh\beta t} = \frac{\alpha\beta}{2(\beta-1)}\cdot\frac{\sinh t\cosh\alpha t\sinh(\beta-1)t}{\sinh\beta t\sinh\alpha t}\\ &= \frac{\alpha\beta}{2(\beta-1)}\cdot\frac{\sinh\frac{1}{\beta}(\beta t)\cosh\frac{\alpha}{\beta}(\beta t)}{\sinh\beta t}\cdot\frac{\sinh(\beta-1)t}{\sinh\alpha t}.\end{aligned}$$

By [26, Proposition 5], the function $\frac{\sinh\frac{1}{\beta}(\beta t)\cosh\frac{\alpha}{\beta}(\beta t)}{\sinh\beta t}$ is positive definite, if $\frac{1}{\beta}+\frac{\alpha}{\beta}\leq 1$ and $\frac{\alpha}{\beta}\leq\frac{1}{2}$. The function $\frac{\sinh(\beta-1)t}{\sinh\alpha t}$ is also positive definite, if $\beta-1\leq\alpha$. From these conditions, we have $\beta=\alpha+1$ and $\alpha\leq 1$. The case $\alpha=\frac{1}{m}$ and $\beta=\frac{m+1}{m}$ satisfies the above conditions. Thus we have $A_{\frac{1}{m}}\preceq M_{\frac{m+1}{m}}$ which leads to the last inequality. □

Remark 4.1.1. Since $\frac{m+1}{m} < \frac{m}{m-1}$, by [111, Theorem 2.1], we have $M_{\frac{m+1}{m}} \preceq M_{\frac{m}{m-1}}$. Thus we have

$$\left\|M_{\frac{m+1}{m}}(S,T)X\right\|_u \le \left\|M_{\frac{m}{m-1}}(S,T)X\right\|_u,$$

which means Theorem 4.1.3 gives a general result for Theorem 4.1.1.

Remark 4.1.2. From the well-known fact $M_\alpha \preceq M_\beta$ for $\alpha < \beta$, we have $H_1 = HM = M_{-1} \preceq M_{1/2} = GM$ and $H_{1/2} = GM = M_{1/2} \preceq M_{2/3}$. Thus we have

$$\left\|H_{\frac{1}{m}}(S,T)X\right\|_u \le \left\|M_{\frac{m}{m+1}}(S,T)X\right\|_u, \tag{4.1.2}$$

for any $S, T, X \in \mathbb{B}(\mathcal{H})$ with $S, T \ge 0$, $m = 1, 2$ and any unitarily invariant norm $\|\cdot\|_u$.

However, we do not have the scalar inequality $H_{1/3}(t,1) \le M_{3/4}(t,1)$ for $t > 0$ in general, so that the inequality (4.1.2) is not true for $m = 3$. We also do not have the scalar inequality $H_{1/3}(t,1) \ge M_{3/4}(t,1)$ for $t > 0$, in general.

We obtained new and tight bounds of the logarithmic mean for a unitarily invariant norm. Our results improved the famous inequalities by F. Hiai and H. Kosaki [110, 111]. Concluding this section, we summarize Theorem 4.1.3 by the familiar form. From the calculations,

$$G_{1/m_1}(s,t) = \frac{1}{m_1}\sum_{k=1}^{m_1} s^{(2k-1)/2m_1} t^{(2m_1-(2k-1))/2m_1}$$

and

$$A_{1/m_2}(s,t) = \frac{1}{m_2}\left(\sum_{k=0}^{m_2} s^{k/m_2} t^{(m_2-k)/m_2} - \frac{1}{2}(s+t)\right),$$

we have

$$G_{1/m_1}(S,T)X = \frac{1}{m_1}\sum_{k=1}^{m_1} S^{(2k-1)/2m_1} X T^{(2m_1-(2k-1))/2m_1}$$

and

$$A_{1/m_2}(S,T)X = \frac{1}{m_2}\left(\sum_{k=0}^{m_2} S^{k/m_2} X T^{(m_2-k)/m_2} - \frac{1}{2}(SX+XT)\right).$$

In addition, from the paper [111], we know that

$$M_{m_1/(m_1+1)}(S,T)X = \frac{1}{m_1}\sum_{k=1}^{m_1} S^{k/(m_1+1)} X T^{(m_1+1-k)/(m_1+1)}$$

and

$$M_{m_2/(m_2-1)}(S,T)X = \frac{1}{m_2}\sum_{k=0}^{m_2-1} S^{k/(m_2-1)} X T^{(m_2-1-k)/(m_2-1)}.$$

Thus Theorem 4.1.3 can be rewritten as the following inequalities which are our main results of the present section:

$$\begin{aligned}
\frac{1}{m_1}\left\|\sum_{k=1}^{m_1} S^{k/(m_1+1)}XT^{(m_1+1-k)/(m_1+1)}\right\|_u &\le \frac{1}{m_1}\left\|\sum_{k=1}^{m_1} S^{(2k-1)/2m_1}XT^{(2m_1-(2k-1))/2m_1}\right\|_u \\
&\le \left\|\int_0^1 S^{\nu}XT^{1-\nu}\,d\nu\right\|_u \\
&\le \frac{1}{m_2}\left\|\sum_{k=0}^{m_2} S^{k/m_2}XT^{(m_2-k)/m_2} - \frac{1}{2}(SX+XT)\right\|_u \\
&\le \frac{1}{m_2}\left\|\sum_{k=0}^{m_2-1} S^{k/(m_2-1)}XT^{(m_2-1-k)/(m_2-1)}\right\|_u, \qquad (4.1.3)
\end{aligned}$$

for $S, T, X \in B(\mathcal{H})$ with $S, T \ge 0$, $m_1 \ge 1$, $m_2 \ge 2$ and any unitarily invariant norm $\|\cdot\|_u$.

Closing this section, we discuss our bounds from the viewpoint of numerical analysis. We rewrite our results by scalar inequalities:

$$\alpha_m(t) \le \int_0^1 t^{\nu}\,d\nu \le \beta_m(t), \quad (m \ge 1, t > 0), \tag{4.1.4}$$

where

$$\alpha_m(t) = \frac{1}{m}\sum_{k=1}^{m} t^{(2k-1)/2m}, \quad \beta_m(t) = \frac{1}{m}\left(\sum_{k=0}^{m} t^{k/m} - \frac{t+1}{2}\right).$$

As we mentioned in the below of Proposition 2.2.1 in Section 2.2, the above inequalities (4.1.4) have been first found as **self-improvement inequalities**. To explain this, we start from the famous inequalities which are the ordering of three means:

$$\sqrt{x} \le \frac{x-1}{\log x}\left(= \int_0^1 x^{\nu}\,d\nu\right) \le \frac{x+1}{2}, \quad (x > 0, x \ne 1). \tag{4.1.5}$$

In (4.1.5) we put $x = \sqrt{t}$, we obtain

$$\sqrt{t} \le \frac{t^{\frac{1}{4}}}{2}(\sqrt{t}+1) \le \frac{t-1}{\log t} \le \left(\frac{\sqrt{t}+1}{2}\right)^2 \le \frac{t+1}{2}.$$

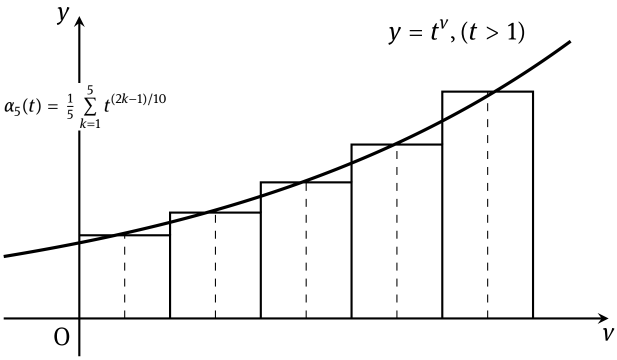
Thus we found the bounds (4.1.4) by putting $x = t^{1/m} > 0$ in (4.1.5). For the above bounds given in (4.1.4), we have the following properties.

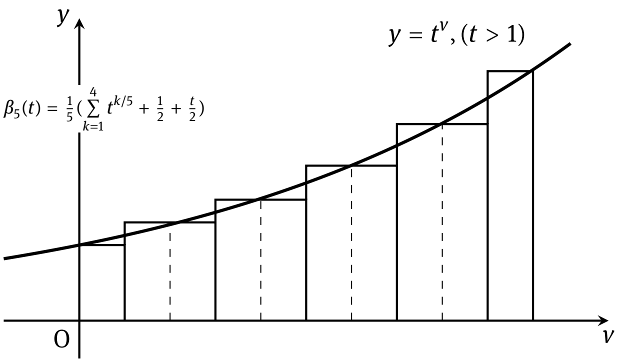
Proposition 4.1.3.
(1) $\alpha_m(t) \le \alpha_{m+1}(t), (t > 0)$.
(2) $\beta_{m+1}(t) \le \beta_m(t), (t > 0)$.

(3) $\lim_{m\to\infty} \alpha_m(t) = L(t,1)$, $\lim_{m\to\infty} \beta_m(t) = L(t,1)$, $(t > 0)$.
(4) $\alpha_m(t) \le \beta_m(t)$, $(t > 0)$,

where $L(t,1) = \frac{t-1}{\log t}$ *is a representing function of the logarithmic mean.*

The examples $\alpha_5(t)$ and $\beta_5(t)$ for the function $y = t^v$, $(t > 1)$ are given in the following figures:





We compare our bounds to standard Riemannian bounds given in

$$\begin{cases} \gamma_m(t) < \int_0^1 t^v\, dv < \delta_m(t), & (m \ge 1, 0 < t < 1), \\ \delta_m(t) < \int_0^1 t^v\, dv < \gamma_m(t), & (m \ge 1, t > 1), \end{cases}$$

where

$$\gamma_m(t) = \frac{1}{m}\sum_{k=1}^{m} t^{k/m}, \quad \delta_m(t) = \frac{1}{m}\sum_{k=0}^{m-1} t^{k/m}.$$

Then we have the following relation.

Proposition 4.1.4.
(1) *If* $0 < t < 1$, *then* $\alpha_m(t) > \gamma_m(t)$ *and* $\beta_m(t) < \delta_m(t)$.
(2) *If* $t > 1$, *then* $\alpha_m(t) > \delta_m(t)$ *and* $\beta_m(t) < \gamma_m(t)$.

Proof. It follows from

$$\alpha_m(t) - \gamma_m(t) = \frac{(1-t)}{m(t^{-\frac{1}{2m}} + 1)} > 0, \quad \beta_m(t) - \delta_m(t) = \frac{t-1}{2m} < 0$$

and

$$\alpha_m(t) - \delta_m(t) = \frac{t-1}{m(t^{\frac{1}{2m}} + 1)} > 0, \quad \beta_m(t) - \gamma_m(t) = \frac{1-t}{2m} < 0. \qquad \square$$

Proposition 4.1.4 shows that our bounds $\alpha_m(t)$ and $\beta_m(t)$ of the logarithmic mean $\int_0^1 t^\nu \, d\nu$ are tighter than the standard Riemannian bounds $\gamma_m(t)$ and $\delta_m(t)$.

4.2 Trace inequality

A trace has cyclic property $\mathrm{Tr}[XYZ] = \mathrm{Tr}[ZXY]$ so that it may have a week noncommutativity since $\mathrm{Tr}[XY] = \mathrm{Tr}[YX]$. A cyclic property will be often useful tool to get the trace inequalities, while it cannot use to get the operator inequalities. For $X, Y, Z > 0$, we have $\mathrm{Tr}[XZYZ] \geq 0$. But we note that for $X, Y > 0$, $\mathrm{Tr}[X] \leq \mathrm{Tr}[Y]$ does not imply

$$\mathrm{Tr}[XZ] \leq \mathrm{Tr}[YZ], \quad \text{for any } Z > 0.$$

Here, we list up some known trace inequalities which are related to this book. Some of them will be applied to obtain our new and/or generalized results in the preceding chapters/sections.

We start from famous **Araki inequality**.

Theorem 4.2.1 ([11]). *Let* $X, Y \geq 0$ *and* $r \geq 1$, $p > 0$. *Then*

$$\mathrm{Tr}[(Y^{1/2}XY^{1/2})^{rp}] \leq \mathrm{Tr}[(Y^{r/2}X^rY^{r/2})^p].$$

Corollary 4.2.1. *Let* $X, Y \geq 0$ *and* $0 \leq r \leq 1$, $p > 0$. *Then*

$$\mathrm{Tr}[(Y^{1/2}XY^{1/2})^{rp}] \geq \mathrm{Tr}[(Y^{r/2}X^rY^{r/2})^p].$$

Proof. Putting $s = rp$ and $t = r(\geq 1)$ in Theorem 4.2.1, we have

$$\mathrm{Tr}[(Y^{1/2}XY^{1/2})^s] \leq \mathrm{Tr}[(Y^{t/2}X^tY^{t/2})^{s/t}], \quad \text{for } t \geq 1,\ s > 0.$$

Furthermore, putting $u = 1/t$ in the above, we have

$$\mathrm{Tr}[(Y^{1/2}XY^{1/2})^s] \leq \mathrm{Tr}[(Y^{1/2u}X^{1/u}Y^{1/2u})^{su}], \quad \text{for } 0 \leq u \leq 1,\ s > 0.$$

Finally, putting $X = A^u$ and $Y = B^u$ in the above, we have

$$\operatorname{Tr}[(B^{u/2}A^uB^{u/2})^s] \le \operatorname{Tr}[(B^{1/2}AB^{1/2})^{su}], \quad \text{for } 0 \le u \le 1, \quad s > 0. \qquad \square$$

From the above inequalities, we have the following.

Corollary 4.2.2. *For $X, Y \ge 0$, we have*
(1) $\operatorname{Tr}[X^pY^p] \le \operatorname{Tr}[(YX^2Y)^{p/2}]$, $0 \le p \le 2$.
(2) $\operatorname{Tr}[X^pY^p] \ge \operatorname{Tr}[(YX^2Y)^{p/2}]$, $p \ge 2$.

The following is known as **Wang–Zhang inequality**.

Theorem 4.2.2 ([232]). *For $X, Y \ge 0$, we have*
(1) $\operatorname{Tr}[(XY)^p] \le \operatorname{Tr}[X^pY^p]$, $|p| \ge 1$.
(2) $\operatorname{Tr}[(XY)^p] \ge \operatorname{Tr}[X^pY^p]$, $|p| \le 1$.

From **Weyl's majorant theorem** ([20, p. 42]), we have

Proposition 4.2.1. *For $X, Y \ge 0$ and $p \ge 0$, we have*

$$\operatorname{Tr}[(XY)^p] \le \operatorname{Tr}[(YX^2Y)^{p/2}].$$

From Corollary 4.2.2, Theorem 4.2.2 and Proposition 4.2.1, we have the following.

Corollary 4.2.3. *For $X, Y \ge 0$, we have*
(1) $\operatorname{Tr}[X^pY^p] \le \operatorname{Tr}[(XY)^p] \le \operatorname{Tr}[(YX^2Y)^{p/2}]$, $0 \le p \le 1$.
(2) $\operatorname{Tr}[(XY)^p] \le \operatorname{Tr}[X^pY^p] \le \operatorname{Tr}[(YX^2Y)^{p/2}]$, $1 \le p \le 2$.
(3) $\operatorname{Tr}[(XY)^p] \le \operatorname{Tr}[(YX^2Y)^{p/2}] \le \operatorname{Tr}[X^pY^p]$, $p \ge 2$.

The following theorem is known as **Lie–Trotter formula**.

Theorem 4.2.3 ([20, Theorem IX.1.3]). *For any two matrices X and Y,*

$$\lim_{p\to 0}(e^{pX}e^{pY})^{1/p} = \lim_{p\to 0}(e^{pY/2}e^{pX}e^{pY/2})^{1/p} = e^{X+Y}.$$

The next theorem follows from Theorem 4.2.1 and Theorem 4.2.3. The following **strengthened Golden–Thompsom inequality** was given by F. Hiai and D. Petz.

Theorem 4.2.4 ([114]). *For two Hermitian matrices X and Y, and $p > 0$,*

$$\operatorname{Tr}[e^{X+Y}] \le \operatorname{Tr}[(e^{pY/2}e^{pX}e^{pY/2})^{1/p}].$$

Definition 4.2.1. For $X, Y \ge 0$, **quantum relative entropy** (so-called **Umegaki entropy** [229]) is defined by

$$D(X|Y) = \operatorname{Tr}[X(\log X - \log Y)].$$

Since the quantum relative entropy is defined by the use of a *trace*, the studies on trace inequalities are often associated with entropy theory in quantum physics. The following relations are known as **variational expressions of relative entropy**.

Lemma 4.2.1 ([201]).
(1) *If A is Hermitian and $Y > 0$, then*

$$\log \mathrm{Tr}[e^{A+\log Y}] = \max\{\mathrm{Tr}[XA] - D(X|Y) : X > 0, \mathrm{Tr}[X] = 1\}.$$

(2) *If $X > 0$, $\mathrm{Tr}[X] = 1$ and B is Hermitian, then*

$$D(X|e^B) = \max\{\mathrm{Tr}[XA] - \log \mathrm{Tr}[e^{A+B}] : A = A^*\}.$$

In the paper [58], a kind of generalization of the variational expressions for the Tsallis relative entropy was given. The following is the logarithmic version of Theorem 4.2.4.

Theorem 4.2.5 ([114]). *For $X, Y \geq 0$, and $p > 0$,*

$$\frac{1}{p}\mathrm{Tr}[X \log Y^{p/2} X^p Y^{p/2}] \leq \mathrm{Tr}[X(\log X + \log Y)]. \tag{4.2.1}$$

The following inequality was given by F. Hiai and D. Petz.

Theorem 4.2.6 ([114]). *For $X, Y \geq 0$, and $p > 0$,*

$$\mathrm{Tr}[X(\log X + \log Y)] \leq \frac{1}{p}\mathrm{Tr}[X \log X^{p/2} Y^p X^{p/2}]. \tag{4.2.2}$$

Remark 4.2.1. The inequality (4.2.1) is equivalent to the following inequality:

$$\mathrm{Tr}[X^{1/p} \log Y^{1/2} X Y^{1/2}] \leq \mathrm{Tr}[X^{1/p}(\log X + \log Y)], \quad p > 0.$$

The inequality (4.2.2) is also equivalent to the following inequality:

$$\mathrm{Tr}[X^{1/p}(\log X + \log Y)] \leq \mathrm{Tr}[X^{1/p} \log X^{1/2} Y X^{1/2}], \quad p > 0.$$

Definition 4.2.2 ([52]). For $X, Y > 0$, the **relative operator entropy** is defined by

$$S(X|Y) = X^{1/2} \log(X^{-1/2} Y X^{-1/2}) X^{1/2}.$$

As a corollary of (4.2.2), we have the relation between quantum relative entropy and relative operator entropy.

Corollary 4.2.4.

$$D(X|Y) \leq -\mathrm{Tr}[S(X|Y)].$$

The following inequality is known as **Marshall–Olkin inequality**.

Theorem 4.2.7 ([152]). *For self-adjoint matrix A and any matrix X, if $g : \mathbb{R} \to \mathbb{R}$ is increasing function, then*

$$\mathrm{Tr}[AX^* g(A)X] \le \mathrm{Tr}[Ag(A)XX^*].$$

We call the following inequalities **Bourin–Fujii inequalities**.

Theorem 4.2.8 ([49]). *For self-adjoint matrices A and X, functions f and g defined on the domain $D \subset \mathbb{R}$, we have*

(i) *If $\{f(a) - f(b)\}\{g(a) - g(b)\} \le 0$ for $a, b \in D$, then*

$$\mathrm{Tr}[f(A)Xg(A)X] \ge \mathrm{Tr}[f(A)g(A)X^2].$$

(ii) *If $\{f(a) - f(b)\}\{g(a) - g(b)\} \ge 0$ for $a, b \in D$, then*

$$\mathrm{Tr}[f(A)Xg(A)X] \le \mathrm{Tr}[f(A)g(A)X^2].$$

4.2.1 Trace inequalities for products of matrices

We start the following proposition in this subsection.

Proposition 4.2.2. *For two real valued functions f, g on $D \subset \mathbb{R}$ and $L \in M_+(n, \mathbb{C})$, $A \in M_h(n, \mathbb{C})$, we have*

$$\mathrm{Tr}[f(A)Lg(A)L] \le \frac{1}{2}\,\mathrm{Tr}[(f(A)L)^2 + (g(A)L)^2]. \tag{4.2.3}$$

Proof. Since the arithmetic mean is greater than the geometric mean and by the use of Schwarz inequality, we have

$$\begin{aligned}\mathrm{Tr}[f(A)Lg(A)L] &= \mathrm{Tr}[(L^{1/2}f(A)L^{1/2})(L^{1/2}g(A)L^{1/2})]\\ &\le (\mathrm{Tr}[(f(A)L)^2])^{1/2}(\mathrm{Tr}[(g(A)L)^2])^{1/2}\\ &\le \frac{1}{2}\,\mathrm{Tr}[(f(A)L)^2 + (g(A)L)^2].\end{aligned}$$ □

Proposition 4.2.3. *For two real valued functions f, g on $D \subset \mathbb{R}$ and $L, A \in M_h(n, \mathbb{C})$, we have*

$$\mathrm{Tr}[f(A)Lg(A)L] \le \frac{1}{2}\,\mathrm{Tr}[f(A)^2L^2 + g(A)^2L^2]. \tag{4.2.4}$$

Proof. Since the arithmetic mean is greater than the geometric mean and by the use of Schwarz inequality, we have

$$\begin{aligned}\mathrm{Tr}[f(A)Lg(A)L] &\le (\mathrm{Tr}[Lf(A)^2L])^{1/2}(\mathrm{Tr}[Lg(A)^2L])^{1/2}\\ &\le \frac{1}{2}\,\mathrm{Tr}[f(A)^2L^2 + g(A)^2L^2].\end{aligned}$$ □

Note that the right-hand side of the inequalities (4.2.3) is less than the right-hand side of the inequalities (4.2.4), since we have $\mathrm{Tr}[XYXY] \le \mathrm{Tr}[X^2Y^2]$ for two Hermitian matrices X and Y in general.

Theorem 4.2.9 ([74]). *For two real valued functions f, g on $D \subset \mathbb{R}$ and $L, A \in M_h(n,\mathbb{C})$, if we have $f(x) \le g(x)$ or $f(x) \ge g(x)$ for any $x \in D$, then we have*

$$\mathrm{Tr}[f(A)Lg(A)L] \le \frac{1}{2}\mathrm{Tr}\big[(f(A)L)^2 + (g(A)L)^2\big]. \tag{4.2.5}$$

Proof. For a spectral decomposition of A such as $A = \sum_k \lambda_k |\phi_k\rangle\langle\phi_k|$, we have

$$\mathrm{Tr}[f(A)Lg(A)L] = \frac{1}{2}\sum_{m,n}\{f(\lambda_m)g(\lambda_n) + g(\lambda_m)f(\lambda_n)\}|\langle\phi_m|L|\phi_n\rangle|^2$$

and

$$\frac{1}{2}\mathrm{Tr}\big[(f(A)L)^2 + (g(A)L)^2\big] = \frac{1}{2}\sum_{m,n}\{f(\lambda_m)f(\lambda_n) + g(\lambda_m)g(\lambda_n)\}|\langle\phi_m|L|\phi_n\rangle|^2.$$

Thus we have the present theorem by

$$\begin{aligned} &f(\lambda_m)f(\lambda_n) + g(\lambda_m)g(\lambda_n) - f(\lambda_m)g(\lambda_n) - g(\lambda_m)f(\lambda_n) \\ &\quad = \{f(\lambda_m) - g(\lambda_m)\}\{f(\lambda_n) - g(\lambda_n)\} \ge 0. \end{aligned}$$ □

Trace inequalities for multiple products of two matrices have been studied by T. Ando, F. Hiai and K. Okubo in [9] with the notion of majorization. Our results in the present subsection are derived by the elementary calculations without the notion of majorization.

Next, we consider the further specialized forms for the products of matrices. We have the following trace inequalities on the products of the power of matrices.

Proposition 4.2.4.
(i) *For any natural number m and $T, A \in M_+(n,\mathbb{C})$, we have the inequality*

$$\mathrm{Tr}\big[(T^{1/m}A)^m\big] \le \mathrm{Tr}[TA^m]. \tag{4.2.6}$$

(ii) *For $\alpha \in [0,1]$, $T \in M_+(n,\mathbb{C})$ and $A \in M_h(n,\mathbb{C})$, we have the inequalities*

$$\mathrm{Tr}\big[(T^{1/2}A)^2\big] \le \mathrm{Tr}[T^\alpha A T^{1-\alpha}A] \le \mathrm{Tr}[TA^2]. \tag{4.2.7}$$

Proof.
(i) Putting $p = m$, $r = 1/m$, $X = A^m$ and $Y = T$ in Araki inequality [11]:

$$\mathrm{Tr}\big[(Y^{r/2}X^rY^{r/2})^p\big] \le \mathrm{Tr}\big[(Y^{1/2}XY^{1/2})^{rp}\big]$$

for $X, Y \in M_+(n,\mathbb{C})$ and $p > 0$, $0 \le r \le 1$, we have the inequality (4.2.6).

(ii) By the use of Theorem 4.2.8, we straightforwardly have

$$\mathrm{Tr}[T^{\alpha}AT^{1-\alpha}A] \le \mathrm{Tr}[TA^2].$$

Again by the use of Theorem 4.2.8, we have

$$\begin{aligned}\mathrm{Tr}[T^{1/2}AT^{1/2}A] &= \mathrm{Tr}[(T^{1/4}AT^{1/4})^2]\\ &\le \mathrm{Tr}[T^{\alpha-1/2}(T^{1/4}AT^{1/4})T^{1/2-\alpha}(T^{1/4}AT^{1/4})]\\ &= \mathrm{Tr}[T^{\alpha}AT^{1-\alpha}A]. \qquad \square\end{aligned}$$

In this subsection, we study a generalization of Proposition 4.2.4 as an application of arithmetic–geometric mean inequality.

Theorem 4.2.10 ([74]). *For positive numbers $p_1, p_2, \dots, p_m$ with $p_1 + p_2 + \cdots + p_m = 1$ and $T, A \in M_+(2, \mathbb{C})$, we have the inequalities*

$$\mathrm{Tr}[(T^{1/m}A)^m] \le \mathrm{Tr}[T^{p_1}AT^{p_2}A \cdots T^{p_m}A] \le \mathrm{Tr}[TA^m]. \tag{4.2.8}$$

Proof. We write $T = \sum_{i=0}^{1} \lambda_i |\psi_i\rangle\langle\psi_i|$, where we can take $\{|\psi_0\rangle, |\psi_1\rangle\}$ as a complete orthonormal base. By the use of the arithmetic–geometric mean inequality with weights $p_1, \dots, p_m$, we then have

$$\begin{aligned}&\mathrm{Tr}[T^{p_1}AT^{p_2}A \cdots T^{p_m}A]\\ &= \sum_{i_1,i_2,\dots,i_m} \left(\frac{\lambda_{i_1}^{p_1}\lambda_{i_2}^{p_2}\cdots\lambda_{i_m}^{p_m} + \lambda_{i_1}^{p_2}\lambda_{i_2}^{p_3}\cdots\lambda_{i_m}^{p_1} + \cdots + \lambda_{i_1}^{p_m}\lambda_{i_2}^{p_1}\cdots\lambda_{i_m}^{p_{m-1}}}{m}\right)\\ &\quad \times \langle\psi_{i_1}|A|\psi_{i_2}\rangle\langle\psi_{i_2}|A|\psi_{i_3}\rangle\cdots\langle\psi_{i_m}|A||\psi_{i_1}\rangle\\ &\le \sum_{i_1,i_2,\dots,i_m} \left(\frac{p_1\lambda_{i_1} + \cdots + p_m\lambda_{i_m} + p_2\lambda_{i_1} + \cdots + p_1\lambda_{i_m} + \cdots + p_m\lambda_{i_1} + \cdots + p_{m-1}\lambda_{i_m}}{m}\right)\\ &\quad \times \langle\psi_{i_1}|A|\psi_{i_2}\rangle\langle\psi_{i_2}|A|\psi_{i_3}\rangle\cdots\langle\psi_{i_m}|A|\psi_{i_1}\rangle\\ &= \sum_{i_1,i_2,\dots,i_m} \left(\frac{\lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_m}}{m}\right)\langle\psi_{i_1}|A|\psi_{i_2}\rangle\langle\psi_{i_2}|A|\psi_{i_3}\rangle\cdots\langle\psi_{i_m}|A|\psi_{i_1}\rangle\\ &= \mathrm{Tr}[TA^m].\end{aligned} \tag{4.2.9}$$

We should note that the above calculation is assured by

$$\langle\psi_{i_1}|A|\psi_{i_2}\rangle\langle\psi_{i_2}|A|\psi_{i_3}\rangle\cdots\langle\psi_{i_m}|A|\psi_{i_1}\rangle \ge 0,$$

since every i_j ($j = 1, 2, \dots, m$) takes 0 or 1. See Lemma 4.2.2 in the below. Again by the use of the arithmetic–geometric mean inequality without weight, we have

$$\mathrm{Tr}[T^{p_1}AT^{p_2}A \cdots T^{p_m}A]$$

$$
\begin{aligned}
&= \sum_{i_1,i_2,\ldots,i_m} \left(\frac{\lambda_{i_1}^{p_1}\lambda_{i_2}^{p_2}\cdots\lambda_{i_m}^{p_m} + \lambda_{i_1}^{p_2}\lambda_{i_2}^{p_3}\cdots\lambda_{i_m}^{p_1} + \cdots + \lambda_{i_1}^{p_m}\lambda_{i_2}^{p_1}\cdots\lambda_{i_m}^{p_{m-1}}}{m} \right) \\
&\quad \times \langle\psi_{i_1}|A|\psi_{i_2}\rangle\langle\psi_{i_2}|A|\psi_{i_3}\rangle\cdots\langle\psi_{i_m}|A|\psi_{i_1}\rangle \\
&\geq \sum_{i_1,i_2,\ldots,i_m} \lambda_{i_1}^{1/m}\lambda_{i_2}^{1/m}\cdots\lambda_{i_m}^{1/m}\langle\psi_{i_1}|A|\psi_{i_2}\rangle\langle\psi_{i_2}|A|\psi_{i_3}\rangle\cdots\langle\psi_{i_m}|A|\psi_{i_1}\rangle \\
&= \mathrm{Tr}\big[(T^{1/m}A)^m\big]. \qquad (4.2.10)
\end{aligned}
$$

Note that the inequalities (4.2.9) and (4.2.10) hold even if $\lambda_0 = 0$ or $\lambda_1 = 0$. It is a trivial when $\lambda_0 = 0$ and $\lambda_1 = 0$. Thus the proof of the present theorem is completed. □

Note that the second inequality of (4.2.8) is derived by putting $f_i(x) = x^{p_i}$ and $g_i(x) = x$ for $i = 1,\ldots,m$ in [9, Theorem 4.1]. However, the first inequality of (4.2.8) cannot be derived by applying [9, Theorem 4.1].

Remark 4.2.2. From the process of the proof in Theorem 4.2.10, we find that, if T is an invertible, then the equalities in both inequalities (4.2.9) and (4.2.10) hold if and only if $T = kI$.

Lemma 4.2.2. *For $A \in M_+(2,\mathbb{C})$ and a complete orthonormal base $\{\psi_0,\psi_1\}$ of $\mathbb{C}^2$, we have*

$$\langle\psi_{i_1}|A|\psi_{i_2}\rangle\langle\psi_{i_2}|A|\psi_{i_3}\rangle\cdots\langle\psi_{i_m}|A|\psi_{i_1}\rangle \geq 0,$$

where every i_j ($j = 1,2,\ldots,m$) takes 0 *or* 1.

Proof. We set a symmetric group by

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & m \\ m & 1 & \cdots & m-1 \end{pmatrix}.$$

We also set

$$S = \{1 \leq j \leq m \mid i_j = i_{\pi(j)}\}$$

for $(i_1,i_2,\ldots,i_m) \in \{0,1\}^m$. Then we have

$$\prod_{j=1}^m \langle\psi_{i_{\pi(j)}}|A|\psi_{i_j}\rangle = \prod_{j\in S}\langle\psi_{i_{\pi(j)}}|A|\psi_{i_j}\rangle \cdot \prod_{j\notin S}\langle\psi_{i_{\pi(j)}}|A|\psi_{i_j}\rangle.$$

Here, we have

$$\prod_{j\in S}\langle\psi_{i_{\pi(j)}}|A|\psi_{i_j}\rangle \geq 0,$$

since A is a nonnegative matrix. In addition, $m - |S|$ necessarily takes 0 or an even number (see Lemma 4.2.3 in the below) and then we have

$$\prod_{j\notin S}\langle\psi_{i_{\pi(j)}}|A|\psi_{i_j}\rangle = \big|\langle\psi_0|A|\psi_1\rangle\big|^{m-|S|} \geq 0.$$

Therefore, we have the present lemma. □

Lemma 4.2.3. *Image the situation that we put arbitrary l_0 vectors $|\psi_0\rangle$ and l_1 vectors $|\psi_1\rangle$ on the circle and then the circle is divided into $l_0 + l_1$ circular arcs. Then the number $|S^c|$ of the circular arcs having different edges is* 0 *or an even number.*

Proof. If $l_0 = 0$ or $l_1 = 0$, then $|S^c| = 0$. Thus we consider $l_0 \neq 0$ and $l_1 \neq 0$. We suppose that the number of the circular arcs having same $|\psi_0\rangle$ in both edges is E, and the number of the circular arcs having $|\psi_0\rangle$ and $|\psi_1\rangle$ in their edges is F. We now form the circular arcs by combining every $|\psi_0\rangle$ with its both sides. (We do not consider the circular arcs formed by the other method.) Thus the number of the circular arcs formed by the above method is an even number, since every $|\psi_0\rangle$ forms two circular arcs. In addition, its number coincides with $2E+F$ which shows the number that every $|\psi_0\rangle$ is doubly counted. Thus $2E + F$ takes an even number. Then F must be an even number. □

We give an example of Lemma 4.2.2 for readers' convenience.

Example 4.2.1. For $\{i_1, i_2, i_3, i_4, i_5, i_6, i_7\} = \{0, 0, 1, 1, 0, 1, 0\}$, we have $S = \{1, 2, 4\}$ and $S^c = \{3, 5, 6, 7\}$. Then we have

$$\prod_{j\in S} \langle \psi_{i_{\pi(j)}} | A | \psi_{i_j} \rangle = \langle \psi_0 | A | \psi_0 \rangle \langle \psi_0 | A | \psi_0 \rangle \langle \psi_1 | A | \psi_1 \rangle \geq 0$$

and

$$\prod_{j\notin S} \langle \psi_{i_{\pi(j)}} | A | \psi_{i_j} \rangle = \langle \psi_0 | A | \psi_1 \rangle \langle \psi_1 | A | \psi_0 \rangle \langle \psi_0 | A | \psi_1 \rangle \langle \psi_1 | A | \psi_0 \rangle = \left| \langle \psi_0 | A | \psi_1 \rangle \right|^4 \geq 0.$$

Thus we have

$$\prod_{j=1}^{7} \langle \psi_{i_{\pi(j)}} | A | \psi_{i_j} \rangle = \prod_{j\in S} \langle \psi_{i_{\pi(j)}} | A | \psi_{i_j} \rangle \cdot \prod_{j\notin S} \langle \psi_{i_{\pi(j)}} | A | \psi_{i_j} \rangle \geq 0.$$

Remark 4.2.3. For $T, A \in M_+(n, \mathbb{C})$, there exist T, A and $p_1, p_2, \ldots, p_m$ such that

$$\mathrm{Tr}\left[\prod_{i=1}^{m} (T^{p_i} A) \right] \notin \mathbb{R},$$

if $n \geq 3$ and $m \geq 3$. See [74] for the example. Therefore, the inequalities (4.2.8) does not make a sense in such more general cases than Theorem 4.2.10.

We still have the following conjectures.

Conjecture 4.2.1. *Do the following inequalities hold or not, for $T, A \in M_+(n, \mathbb{C})$ and positive numbers $p_1, p_2, \ldots, p_m$ with $p_1 + p_2 + \cdots + p_m = 1$?*

(i) $\mathrm{Tr}[(T^{1/m}A)^m] \leq \mathrm{Re}\{\mathrm{Tr}[T^{p_1} A T^{p_2} A \cdots T^{p_m} A]\}$.

(ii) $|\mathrm{Tr}[T^{p_1} A T^{p_2} A \cdots T^{p_m} A]| \leq \mathrm{Tr}[TA^m]$.

4.2.2 Conjectured trace inequality

The purpose of this subsection is to give the answer to the following conjecture which was given in the paper [59].

Conjecture 4.2.2 ([59]). *For $X, Y \in M_+(n, \mathbb{C})$ and $p \in \mathbb{R}$, the following inequalities hold or not:*
(i) $\mathrm{Tr}[(I + X + Y + Y^{1/2}XY^{1/2})^p] \le \mathrm{Tr}[(I + X + Y + XY)^p]$ *for* $p \ge 1$.
(ii) $\mathrm{Tr}[(I + X + Y + Y^{1/2}XY^{1/2})^p] \ge \mathrm{Tr}[(I + X + Y + XY)^p]$ *for* $0 \le p \le 1$.

We first note that the matrix $I + X + Y + XY = (I + X)(I + Y)$ is generally not positive semi-definite. However, the eigenvalues of the matrix $(I + X)(I + Y)$ are same to those of the positive semi-definite matrix $(I + X)^{1/2}(I + Y)(I + X)^{1/2}$. Therefore, the expression $\mathrm{Tr}[(I + X + Y + XY)^p]$ always makes sense.

We easily find that the equality for (i) and (ii) in Conjecture 4.2.2 holds in the case of $p = 1$. In addition, the case of $p = 2$ was proven by elementary calculations in [59].

Putting $T = (I + X)^{1/2}$ and $S = Y^{1/2}$, Conjecture 4.2.2 can be reformulated by the following problem, because we have $\mathrm{Tr}[(I+X+Y+XY)^p] = \mathrm{Tr}[(T^2+T^2S^2)^p] = \mathrm{Tr}[(T^2(I+S^2))^p] = \mathrm{Tr}[(T(I + S^2)T)^p] = \mathrm{Tr}[(T^2 + TS^2T)^p]$.

Problem 4.2.1. For $T, S \in M_+(n, \mathbb{C})$ and $p \in \mathbb{R}$, the following inequalities hold or not:
(i) $\mathrm{Tr}[(T^2 + ST^2S)^p] \le \mathrm{Tr}[(T^2 + TS^2T)^p]$ for $p \ge 1$.
(ii) $\mathrm{Tr}[(T^2 + ST^2S)^p] \ge \mathrm{Tr}[(T^2 + TS^2T)^p]$ for $0 \le p \le 1$.

To solve Problem 4.2.1, we use the concept of the **majorization**. See [151] for the details on the majorization. Here, for $X \in M_h(n, \mathbb{C})$, $\lambda^\downarrow(X) = (\lambda_1^\downarrow(X), \dots, \lambda_n^\downarrow(X))$ represents the eigenvalues of the Hermitian matrix X in decreasing order, $\lambda_1^\downarrow(X) \ge \cdots \ge \lambda_n^\downarrow(X)$. In addition, $x \prec y$ means that $x = (x_1, \dots, x_n)$ is majorized by $y = (y_1, \dots, y_n)$, if we have

$$\sum_{j=1}^{k} x_j \le \sum_{j=1}^{k} y_j, \quad (k = 1, \dots, n-1), \qquad \sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j.$$

We need the following lemma which can be obtained as a consequence of **Ky Fan's maximum principle**.

Lemma 4.2.4 ([20, p. 35]). *For $A, B \in M_h(n, \mathbb{C})$ and any $k = 1, 2, \dots, n$, we have*

$$\sum_{j=1}^{k} \lambda_j^\downarrow(A + B) \le \sum_{j=1}^{k} \lambda_j^\downarrow(A) + \sum_{j=1}^{k} \lambda_j^\downarrow(B). \tag{4.2.11}$$

Then we have the following theorem.

Theorem 4.2.11 ([75]). *For $S, T \in M_+(n, \mathbb{C})$, we have*

$$\lambda^\downarrow(T^2 + ST^2S) \prec \lambda^\downarrow(T^2 + TS^2T). \tag{4.2.12}$$

Proof. For $S, T \in M_+(n, \mathbb{C})$, we need only to show the following:

$$\sum_{j=1}^{k} \lambda_j^{\downarrow}(T^2 + ST^2S) \le \sum_{j=1}^{k} \lambda_j^{\downarrow}(T^2 + TS^2T) \tag{4.2.13}$$

for $k = 1, 2, \ldots, n-1$, since we have

$$\sum_{j=1}^{n} \lambda_j^{\downarrow}(T^2 + ST^2S) = \sum_{j=1}^{n} \lambda_j^{\downarrow}(T^2 + TS^2T),$$

which is equivalent to $\mathrm{Tr}[T^2 + ST^2S] = \mathrm{Tr}[T^2 + TS^2T]$. By Lemma 4.2.4, we have

$$2\sum_{j=1}^{k} \lambda_j^{\downarrow}(X) \le \sum_{j=1}^{k} \lambda_j^{\downarrow}(X+Y) + \sum_{j=1}^{k} \lambda_j^{\downarrow}(X-Y), \tag{4.2.14}$$

for $X, Y \in M_h(n, \mathbb{C})$ and any $k = 1, 2, \ldots, n$. For $X \in M(n, \mathbb{C})$, the matrices XX^* and X^*X are unitarily similar so that we have $\lambda_j^{\downarrow}(XX^*) = \lambda_j^{\downarrow}(X^*X)$. Then we have the following inequality:

$$\begin{aligned}
2\sum_{j=1}^{k} \lambda_j^{\downarrow}(T^2 + TS^2T) &= \sum_{j=1}^{k} \lambda_j^{\downarrow}(T^2 + TS^2T) + \sum_{j=1}^{k} \lambda_j^{\downarrow}(T^2 + TS^2T) \\
&= \sum_{j=1}^{k} \lambda_j^{\downarrow}((T + iTS)(T - iST)) + \sum_{j=1}^{k} \lambda_j^{\downarrow}((T - iTS)(T + iST)) \\
&= \sum_{j=1}^{k} \lambda_j^{\downarrow}((T - iST)(T + iTS)) + \sum_{j=1}^{k} \lambda_j^{\downarrow}((T + iST)(T - iTS)) \\
&= \sum_{j=1}^{k} \lambda_j^{\downarrow}(T^2 + ST^2S + i(T^2S - ST^2)) + \sum_{j=1}^{k} \lambda_j^{\downarrow}(T^2 + ST^2S - i(T^2S - ST^2)) \\
&\ge 2\sum_{j=1}^{k} \lambda_j^{\downarrow}(T^2 + ST^2S),
\end{aligned}$$

for any $k = 1, 2, \ldots, n-1$, by using the inequality (4.2.14) for $X = T^2 + ST^2S$ and $Y = i(T^2S - ST^2)$. Thus we have the inequality (4.2.13) so that the proof is completed. □

From Theorem 4.2.11, we have the following corollary.

Corollary 4.2.5. *For $T, S \in M_+(n, \mathbb{C})$ and $p \in \mathbb{R}$, the following inequalities hold:*
(i) $\mathrm{Tr}[(T^2 + ST^2S)^p] \le \mathrm{Tr}[(T^2 + TS^2T)^p]$ *for* $p \ge 1$.
(ii) $\mathrm{Tr}[(T^2 + ST^2S)^p] \ge \mathrm{Tr}[(T^2 + TS^2T)^p]$ *for* $0 \le p \le 1$.

Proof. Since $f(x) = x^p$, $(p \ge 1)$ is convex function and $f(x) = x^p$, $(0 \le p \le 1)$ is concave function, we have the present corollary thanks to Theorem 4.2.11 and a general property of majorization (see in [20, p. 40]). □

As mentioned in the beginning of this subsection, Corollary 4.2.5 implies the following corollary by putting $T = (I+X)^{1/2}$ and $S = Y^{1/2}$.

Corollary 4.2.6. *For $X, Y \in M_+(n,\mathbb{C})$ and $p \in \mathbb{R}$, the following inequalities hold:*
(i) $\mathrm{Tr}[(I+X+Y+Y^{1/2}XY^{1/2})^p] \le \mathrm{Tr}[(I+X+Y+XY)^p]$ *for* $p \ge 1$.
(ii) $\mathrm{Tr}[(I+X+Y+Y^{1/2}XY^{1/2})^p] \ge \mathrm{Tr}[(I+X+Y+XY)^p]$ *for* $0 \le p \le 1$.

Thus Conjecture 4.2.2 was completely solved with an affirmative answer. After our solution, the generalized results have been shown later. We give them here.

Theorem 4.2.12 ([227]). *For $A, B \in M_+(n,\mathbb{C})$,*

$$2\lambda^{\downarrow}(AA^* + BB^*) \prec \lambda^{\downarrow}(A^*A + B^*B - C) + \lambda^{\downarrow}(A^*A + B^*B + C),$$

where $C = A^*B + B^*A$.

Proof. Since $\lambda^{\downarrow}(H+K) \prec \lambda^{\downarrow}(H) + \lambda^{\downarrow}(K), \lambda^{\downarrow}(XY) = \lambda^{\downarrow}(YX)$ in general, we have

$$\begin{aligned}
&2\lambda^{\downarrow}(AA^* + BB^*)\\
&= \lambda^{\downarrow}\left([A,B]\begin{bmatrix} I & -I \\ -I & I \end{bmatrix}\begin{bmatrix} A^* \\ B^* \end{bmatrix} + [A,B]\begin{bmatrix} I & I \\ I & I \end{bmatrix}\begin{bmatrix} A^* \\ B^* \end{bmatrix}\right)\\
&\prec \lambda^{\downarrow}\left([A,B]\begin{bmatrix} I & -I \\ -I & I \end{bmatrix}\begin{bmatrix} A^* \\ B^* \end{bmatrix}\right) + \lambda^{\downarrow}\left([A,B]\begin{bmatrix} I & I \\ I & I \end{bmatrix}\begin{bmatrix} A^* \\ B^* \end{bmatrix}\right)\\
&= \lambda^{\downarrow}\left([A,B]\begin{bmatrix} I \\ -I \end{bmatrix}\cdot[I,-I]\begin{bmatrix} A^* \\ B^* \end{bmatrix}\right) + \lambda^{\downarrow}\left([A,B]\begin{bmatrix} I \\ I \end{bmatrix}\cdot[I,I]\begin{bmatrix} A^* \\ B^* \end{bmatrix}\right)\\
&= \lambda^{\downarrow}\left([I,-I]\begin{bmatrix} A^* \\ B^* \end{bmatrix}\cdot[A,B]\begin{bmatrix} I \\ -I \end{bmatrix}\right) + \lambda^{\downarrow}\left([I,I]\begin{bmatrix} A^* \\ B^* \end{bmatrix}\cdot[A,B]\begin{bmatrix} I \\ I \end{bmatrix}\right)\\
&= \lambda^{\downarrow}(A^*A + B^*B - A^*B + B^*A) + \lambda^{\downarrow}(A^*A + B^*B + A^*B + B^*A).
\end{aligned}$$ □

Corollary 4.2.7 ([144]). *For $X, Y \in M_+(n,\mathbb{C})$ such that X^*Y is Hermitian or skew-Hermitian,*

$$\lambda^{\downarrow}(XX^* + YY^*) \prec \lambda^{\downarrow}(X^*X + Y^*Y).$$

Proof. If X^*Y is skew-Hermitian, then $C = X^*Y + Y^*X = 0$. If X^*Y is Hermitian, then $X^*(iY)$ is skew-Hermitian so that we replace Y by iY. Thus Theorem 4.2.12 implies Corollary 4.2.7. □

If we take $X = T$ and $Y = ST$ in Corollary 4.2.7, we can recover Theorem 4.2.11. Closing this subsection, we give the proofs for Theorem 4.2.11 for special cases. Actually, at first we could not prove Theorem 4.2.11 for general case directly. First, we tried and succeeded to give proofs for the special cases. Giving notes for such trials may be helpful to the readers.

Proposition 4.2.5 (Special case of Corollary 4.2.5). *For $T, S \in M_+(2,\mathbb{C})$ and $p \in \mathbb{R}$, the following inequalities hold:*
(i) $\mathrm{Tr}[(T^2 + ST^2S)^p] \le \mathrm{Tr}[(T^2 + TS^2T)^p]$ *for* $p \ge 1$.
(ii) $\mathrm{Tr}[(T^2 + ST^2S)^p] \ge \mathrm{Tr}[(T^2 + TS^2T)^p]$ *for* $0 \le p \le 1$.

To prove Proposition 4.2.5, we prepare the following lemma.

Lemma 4.2.5. *For real numbers α, β, γ such that $\alpha > \beta \ge \gamma \ge 0$, we have the following inequalities:*
(i) *If $p \ge 1$, then we have $(\alpha+\beta)^p + (\alpha-\beta)^p \ge (\alpha+\gamma)^p + (\alpha-\gamma)^p$.*
(ii) *If $0 \le p \le 1$, then we have $(\alpha+\beta)^p + (\alpha-\beta)^p \le (\alpha+\gamma)^p + (\alpha-\gamma)^p$.*

Proof. We put $f(\beta) = (\alpha+\beta)^p + (\alpha-\beta)^p - (\alpha+\gamma)^p - (\alpha-\gamma)^p$, $(\alpha > \beta \ge \gamma \ge 0)$. Since $f'(\beta) = p\{(\alpha+\beta)^{p-1} - (\alpha-\beta)^{p-1}\} \ge 0$ for $p \ge 1$, we have $f(x) \ge f(\gamma) = 0$, which proves (i). Since $f'(\beta) = p\{(\alpha+\beta)^{p-1} - (\alpha-\beta)^{p-1}\} \le 0$ for $0 \le p \le 1$, we have $f(x) \le f(\gamma) = 0$, which proves (ii). □

Proof of Proposition 4.2.5. Without loss of generality, we may put

$$S = \begin{pmatrix} x & z \\ \bar{z} & y \end{pmatrix}, \quad T = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad (a, b, x, y > 0, xy > |z|^2 > 0).$$

Then two positive matrices $T^2 + TS^2T$ and $T^2 + ST^2S$ have the following eigenvalues, respectively,

$$\mu_\pm = \frac{m \pm \sqrt{m^2 - l_1}}{2}, \quad \nu_\pm = \frac{m \pm \sqrt{m^2 - l_2}}{2},$$

where

$$\begin{aligned} m &= a^2(1 + x^2 + |z|^2) + b^2(1 + y^2 + |z|^2), \\ l_1 &= 4\{a^2b^2(1+x^2)(1+y^2) + 2a^2b^2(1-xy)|z|^2 + a^2b^2|z|^4\}, \\ l_2 &= 4\{a^2b^2(1+x^2)(1+y^2) + (a^4 + b^4 - 2a^2b^2xy)|z|^2 + a^2b^2|z|^4\}. \end{aligned}$$

Then we have

$$\begin{aligned} &\mathrm{Tr}\big[(T^2 + TS^2T)^p\big] - \mathrm{Tr}\big[(T^2 + ST^2S)^p\big] \\ &= \mu_+^p + \mu_-^p - \nu_+^p - \nu_-^p \\ &= \left(\frac{m + \sqrt{m^2 - l_1}}{2}\right)^p + \left(\frac{m - \sqrt{m^2 - l_1}}{2}\right)^p - \left(\frac{m + \sqrt{m^2 - l_2}}{2}\right)^p - \left(\frac{m - \sqrt{m^2 - l_2}}{2}\right)^p. \end{aligned}$$

We easily find $l_2 - l_1 = 4(a^2 - b^2)^2|z|^2 \ge 0$. We also find that $l_1 > 0$. Indeed, $\frac{l_1}{4a^2b^2} = t^2 + 2(1-xy)t + (1+x^2)(1+y^2) > 0$, since $(1-xy)^2 - (1+x^2)(1+y^2) = -(x+y)^2 < 0$,

where we put $t = |z|^2 > 0$. Thus $l_2 \ge l_1 > 0$ implies $m > \sqrt{m^2 - l_1} \ge \sqrt{m^2 - l_2} \ge 0$. The last inequality is caused that matrix $T^2 + ST^2S$ is a Hermitian. We thus have the present theorem, applying Lemma 4.2.5 with $\alpha = \frac{m}{2}$, $\beta = \frac{\sqrt{m^2-l_1}}{2}$ and $\gamma = \frac{\sqrt{m^2-l_2}}{2}$. □

Thus we proved that Conjecture 4.2.2 is true for 2×2 positive matrices. Second, we shall prove that Conjecture 4.2.2 is true for 3×3 positive matrices.

Proposition 4.2.6 (Special case of Theorem 4.2.11). *For $S, T \in M_+(3, \mathbb{C})$, we have*

$$\lambda^{\downarrow}(T^2 + ST^2S) \prec \lambda^{\downarrow}(T^2 + TS^2T). \tag{4.2.15}$$

Proof. For a Hermitian matrix X, we assume that its eigenvalues are arranged by $\lambda_1(X) \ge \lambda_2(X) \ge \lambda_3(X)$. For $S, T \in M_+(3, \mathbb{C})$, we need only to show the following:

$$\lambda_1(T^2 + ST^2S) \le \lambda_1(T^2 + TS^2T) \tag{4.2.16}$$

and

$$\lambda_3(T^2 + ST^2S) \ge \lambda_3(T^2 + TS^2T) \tag{4.2.17}$$

since

$$\sum_{j=1}^{3} \lambda_j(T^2 + ST^2S) = \sum_{j=1}^{3} \lambda_j(T^2 + TS^2T).$$

We may assume the invertibility of T and S. To prove (4.2.16), it is sufficient to prove that $\lambda_1(T^2 + TS^2T) \le 1$ implies $\lambda_1(T^2 + ST^2S) \le 1$, because the both sides of (4.2.16) have the same order of homogeneity for T and S, so that we can multiply T and S by a positive constant.

From $\lambda_1(T^2 + TS^2T) \le 1$, we have $T(I + S^2)T \le I$, namely $I + S^2 \le T^{-2}$. If we take the inverse, then we have $T^2 \le (I + S^2)^{-1}$. Then we have

$$ST^2S \le S(I + S^2)^{-1}S = S\{S^{-2} - S^{-2}(S^{-2} + I)^{-1}S^{-2}\}S = I - S^{-1}(S^{-2} + I)^{-1}S^{-1}.$$

Thus we have

$$T^2 + ST^2S \le I - S^{-1}(S^{-2} + I)^{-1}S^{-1} + T^2 \le I. \tag{4.2.18}$$

Indeed, from $I + S^2 \le T^{-2}$, we have $T^{-2} - I \ge S^2$ which implies $S^{-1}(T^{-2} - I)S^{-1} \ge I$. Therefore, we have $S^{-1}T^{-2}S^{-1} \ge S^{-2} + I$. Taking the inverse, we have $ST^2S \le (S^{-2} + I)^{-1}$ which implies $T^2 \le S^{-1}(S^{-2}+I)^{-1}S^{-1}$. Thus we have the last inequality (4.2.18) and then we have $\lambda_1(T^2 + ST^2S) \le 1$, which completes the proof of (4.2.16).

To prove (4.2.17), by the similar argument of the above, it is sufficient to prove that $\lambda_3(T^2 + TS^2T) \ge 1$ implies $\lambda_3(T^2 + ST^2S) \ge 1$. From $\lambda_3(T^2 + TS^2T) \ge 1$, we have

$T(I+S^2)T \geq I$, namely $I+S^2 \geq T^{-2}$ which implies $T^2 \geq (I+S^2)^{-1}$ by taking the inverse. Then we have

$$ST^2S \geq S(I+S^2)^{-1}S = S\{S^{-2} - S^{-2}(I+S^{-2})^{-1}S^{-2}\}S = I - S^{-1}(I+S^{-2})^{-1}S^{-1}.$$

Thus we have

$$T^2 + ST^2S \geq I + T^2 - S^{-1}(I+S^{-2})^{-1}S^{-1} \geq I. \tag{4.2.19}$$

Indeed, from $T^{-2} \leq S^2 + I$, we have $T^{-2} - I \leq S^2$ which implies $S^{-1}(T^{-2} - I)S^{-1} \leq I$. Therefore we have $S^{-1}T^{-2}S^{-1} \leq S^{-2} + I$. Taking the inverse of the both sides, we have $ST^2S \geq (S^{-2}+I)^{-1}$ which implies $T^2 \geq S^{-1}(S^{-2}+I)^{-1}S^{-1}$. Thus we have the last inequality (4.2.19) and then we have $\lambda_3(T^2 + ST^2S) \geq 1$, which completes the proof of (4.2.17). □

4.2.3 Belmega–Lasaulce–Debbah inequality

E.-V. Belmega, S. Lasaulce and M. Debbah obtained the following trace inequality for positive definite matrices.

Theorem 4.2.13 ([18]). *For $A, B \in M_+(n,\mathbb{C})$ with invertible, and $C, D \in M_+(n,\mathbb{C})$ we have*

$$\mathrm{Tr}\big[(A-B)(B^{-1}-A^{-1}) + (C-D)\{(B+D)^{-1} - (A+C)^{-1}\}\big] \geq 0. \tag{4.2.20}$$

We call the inequality (4.2.20) **Belmega–Lasaulce–Debbah inequality**. In this subsection, we first prove a certain trace inequality for products of matrices, and then as its application, we give a simple proof of (4.2.20). At the same time, our alternative proof gives a refinement of Theorem 4.2.13. In this subsection, we prove the following theorem.

Theorem 4.2.14 ([76]). *For $A, B \in M_+(n,\mathbb{C})$ with invertible, and $C, D \in M_+(n,\mathbb{C})$, we have*

$$\begin{aligned}&\mathrm{Tr}\big[(A-B)(B^{-1}-A^{-1}) + (C-D)\{(B+D)^{-1} - (A+C)^{-1}\}\big]\\ &\quad\geq \big|\mathrm{Tr}\big[(C-D)(B+D)^{-1}(A-B)(A+C)^{-1}\big]\big|.\end{aligned} \tag{4.2.21}$$

To prove this theorem, we need a few lemmas.

Lemma 4.2.6 ([18]). *For $A, B \in M_+(n,\mathbb{C})$ with invertible, $C, D \in M_+(n,\mathbb{C})$ and $X \in M_h(n,\mathbb{C})$, we have*

$$\mathrm{Tr}\big[XA^{-1}XB^{-1}\big] \geq \mathrm{Tr}\big[X(A+C)^{-1}X(B+D)^{-1}\big].$$

Lemma 4.2.7. *For $X, Y \in M(n,\mathbb{C})$, we have*

$$\mathrm{Tr}[X^*X] + \mathrm{Tr}[Y^*Y] \geq 2\big|\mathrm{Tr}[X^*Y]\big|.$$

Proof. Since $\mathrm{Tr}[X^*X] \geq 0$, by the fact that the arithmetical mean is greater than the geometrical mean and Cauchy–Schwarz inequality, we have

$$\frac{\mathrm{Tr}[X^*X] + \mathrm{Tr}[Y^*Y]}{2} \geq \sqrt{\mathrm{Tr}[X^*X]\,\mathrm{Tr}[Y^*Y]} \geq \left|\mathrm{Tr}[X^*Y]\right|. \qquad \square$$

Theorem 4.2.15 ([76]). *For $X_1, X_2 \in M_h(n,\mathbb{C})$ and $S_1, S_2 \in M_+(n,\mathbb{C})$, we have*

$$\mathrm{Tr}[X_1S_1X_1S_2] + \mathrm{Tr}[X_2S_1X_2S_2] \geq 2\left|\mathrm{Tr}[X_1S_1X_2S_2]\right|.$$

Proof. Applying Lemma 4.2.7, we have

$$\begin{aligned}
&\mathrm{Tr}[X_1S_1X_1S_2] + \mathrm{Tr}[X_2S_1X_2S_2] \\
&\quad = \mathrm{Tr}\big[(S_2^{1/2}X_1S_1^{1/2})(S_1^{1/2}X_1S_2^{1/2})\big] + \mathrm{Tr}\big[(S_2^{1/2}X_2S_1^{1/2})(S_1^{1/2}X_2S_2^{1/2})\big] \\
&\quad \geq 2\left|\mathrm{Tr}\big[(S_2^{1/2}X_1S_1^{1/2})(S_1^{1/2}X_2S_2^{1/2})\big]\right| = 2\left|\mathrm{Tr}[X_1S_1X_2S_2]\right|. \qquad \square
\end{aligned}$$

Remark 4.2.4. Theorem 4.2.15 can be regarded as a kind of the generalization of Proposition 4.2.2 given in [74, Proposition 1.1].

Proof of Theorem 4.2.14. By Lemma 4.2.6, we have

$$\begin{aligned}
&\mathrm{Tr}\big[(A-B)(B^{-1}-A^{-1})\big] \\
&\quad = \mathrm{Tr}\big[(A-B)B^{-1}(A-B)A^{-1}\big] \geq \mathrm{Tr}\big[(A-B)(A+C)^{-1}(A-B)(B+D)^{-1}\big] \\
&\quad = \mathrm{Tr}\big[(A-B)(B+D)^{-1}(A-B)(A+C)^{-1}\big].
\end{aligned}$$

Thus the left-hand side of the inequality (4.2.21) can be bounded from below:

$$\begin{aligned}
&\mathrm{Tr}\big[(A-B)(B^{-1}-A^{-1}) + (C-D)\{(B+D)^{-1} - (A+C)^{-1}\}\big] \\
&\quad \geq \mathrm{Tr}\big[(A-B)(B+D)^{-1}(A-B)(A+C)^{-1} + (C-D)(B+D)^{-1}(C-D)(A+C)^{-1}\big] \\
&\qquad + \mathrm{Tr}\big[(C-D)(B+D)^{-1}(A-B)(A+C)^{-1}\big] \\
&\quad \geq 2\left|\mathrm{Tr}\big[(C-D)(B+D)^{-1}(A-B)(A+C)^{-1}\big]\right| \\
&\qquad + \mathrm{Tr}\big[(C-D)(B+D)^{-1}(A-B)(A+C)^{-1}\big]. \qquad (4.2.22)
\end{aligned}$$

Throughout the process of the above, Lemma 4.2.15 was used in the second inequality. Since we have the following equation:

$$\begin{aligned}
&\mathrm{Tr}\big[(C-D)(B+D)^{-1}(A-B)(A+C)^{-1}\big] \\
&\quad = \mathrm{Tr}\big[(C-D)(B+D)^{-1}\big] - \mathrm{Tr}\big[(C-D)(A+C)^{-1}\big] \\
&\qquad - \mathrm{Tr}\big[(C-D)(B+D)^{-1}(C-D)(A+C)^{-1}\big]
\end{aligned}$$

we have $\mathrm{Tr}[(C-D)(B+D)^{-1}(A-B)(A+C)^{-1}] \in \mathbb{R}$. Therefore, we have

$$(4.2.22) \geq \left|\mathrm{Tr}\big[(C-D)(B+D)^{-1}(A-B)(A+C)^{-1}\big]\right|. \qquad \square$$

An alternative refinement for the Belmega–Lasaulce–Debbah inequality was given in the following.

Theorem 4.2.16 ([76]). *For $A, B \in M_+(n, \mathbb{C})$ with invertible, and $C, D \in M_+(n, \mathbb{C})$, we have*

$$\mathrm{Tr}\left[(A-B)(B^{-1}-A^{-1}) + 4(C-D)\{(B+D)^{-1} - (A+C)^{-1}\}\right] \geq 0.$$

We omit the proof of Theorem 4.2.16 since it can be proven by the similar way to Lemma 4.2.7 and Theorem 4.2.15 with slight modifications.

Remark 4.2.5. We always have $\mathrm{Tr}[(A-B)(B^{-1}-A^{-1})] \geq 0$, however the trace

$$\mathrm{Tr}\left[(C-D)\{(B+D)^{-1} - (A+C)^{-1}\}\right]$$

has possibility to take a negative value. Therefore, Theorem 4.2.16 gives a refinement of Theorem 4.2.13.

By putting $A = rA_1$ and $B = rB_1$ for $A, B \in M_+(n, \mathbb{C})$ with invertible in Theorem 4.2.16, we obtain the following corollary.

Corollary 4.2.8. *For $A, B \in M_+(n, \mathbb{C})$ with invertible, $C, D \in M_+(n, \mathbb{C})$ and $r > 0$, we have*

$$\mathrm{Tr}\left[(A-B)(B^{-1}-A^{-1}) + 4(C-D)\{(rB+D)^{-1} - (rA+C)^{-1}\}\right] \geq 0. \tag{4.2.23}$$

Remark 4.2.6. In the case of $r = 2$ in Corollary 4.2.8, the inequality (4.2.23) corresponds to the scalar inequality:

$$(\alpha-\beta)\left(\frac{1}{4\beta} - \frac{1}{4\alpha}\right) + (\gamma-\delta)\left(\frac{1}{2\beta+\delta} - \frac{1}{2\alpha+\gamma}\right) \geq 0$$

for $\alpha, \beta > 0$ and $\gamma, \delta \geq 0$.

We note that the Belmega–Lasaulce–Debbah inequality was generalized in [19].

Theorem 4.2.17 ([19]). *For $A_1, B_1 \in M_+(n, \mathbb{C})$ with invertible, and $A_k, B_k \in M_+(n, \mathbb{C})$ for $k = 2, \ldots, K$, then the following inequality holds:*

$$\mathrm{Tr}\left[\sum_{k=1}^{K}(A_k - B_k)\left\{\left(\sum_{l=1}^{k} B_l\right)^{-1} - \left(\sum_{l=1}^{k} A_l\right)^{-1}\right\}\right] \geq 0.$$

It may be of interests to consider the refinements given in Theorem 4.2.14 and 4.2.16 for Theorem 4.2.13 can be applied to Theorem 4.2.17.

5 Convex functions and Jensen inequality

Recall that if f is a real continuous function, then the so-called **perspective** [197] of f denoted by $\mathcal{P}_f(A|B)$, is defined as

$$\mathcal{P}_f(A|B) = A^{\frac{1}{2}} f(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}) A^{\frac{1}{2}}, \tag{5.0.1}$$

for $A > 0$ and $B \geq 0$. It is also called **solidarity** [50] for an operator monotone function f. If f is *operator convex*, then some important properties have been proven in [45, 46]. We do not treat a perspective deeply in this book. If we take $f(t) = \log t$ or $f(t) = \ln_r t = \frac{t^r - 1}{r}$, $(t > 0, r \neq 0)$ which was defined in Section 2.4 as r-logarithmic function, then the perspective(solidarity) recovers the relative operator entropy [52] or the Tsallis relative operator entropy [240]. See Chapter 7 for the results on them. If the function $f : J \subseteq \mathbb{R} \to \mathbb{R}$ is *convex*, then the so-called **Jensen inequality**

$$f\left(\sum_{i=1}^{n} w_i x_i\right) \leq \sum_{i=1}^{n} w_i f(x_i), \tag{5.0.2}$$

holds for some positive numbers $w_1, \ldots, w_n$ with $\sum\limits_{i=1}^{n} w_i = 1$ and $x_i \in J$.

Let $f : J \to \mathbb{R}$ be a *convex* function and $x_1, \ldots, x_n \in J$ and $w_1, \ldots, w_n$ positive numbers with $W_n = \sum\limits_{i=1}^{n} w_i$. The famous Jensen inequality asserts that

$$f\left(\frac{1}{W_n} \sum_{i=1}^{n} w_i x_i\right) \leq \frac{1}{W_n} \sum_{i=1}^{n} w_i f(x_i). \tag{5.0.3}$$

In the paper [168], one can find that if $f : J \to \mathbb{R}$ is an *operator convex* function, then

$$f\left(\frac{1}{W_n} \sum_{i=1}^{n} w_i A_i\right) \leq \frac{1}{W_n} \sum_{i=1}^{n} w_i f(A_i) \tag{5.0.4}$$

whenever $A_1, \ldots, A_n$ are self-adjoint operators with spectra contained in J.

The celebrated **Choi–Davis–Jensen inequality** (C-D-J inequality for short) [34, 38] asserts that if $f : J \to \mathbb{R}$ is an *operator convex* and $\Phi : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ is a normalized positive linear map, and A is a self-adjoint operator with spectra contained in J, then

$$f(\Phi(A)) \leq \Phi(f(A)). \tag{5.0.5}$$

5.1 A generalized operator Jensen inequality

The following result that provides an *operator version for the Jensen inequality* is due to B. Mond and J. Pečarić.

https://doi.org/10.1515/9783110643473-005

Theorem 5.1.1 (Operator Jensen inequality for convex function [169]). *. Let $A \in \mathbb{B}(\mathcal{H})$ be a self-adjoint operator with $Sp(A) \subseteq [m, M]$ for some scalars $m < M$. If $f(t)$ is a convex function on $[m, M]$, then*

$$f(\langle Ax, x\rangle) \le \langle f(A)x, x\rangle, \tag{5.1.1}$$

for every unit vector $x \in \mathcal{H}$.

The aim of this section is to find an inequality which contains (5.1.1) as a special case. We start by sorting out some of the notions we will consider. Given a continuous function $f : J \to \mathbb{R}$ defined on the compact interval $J \subset \mathbb{R}$, consider a function $\varphi : J \times \mathbb{R} \to \mathbb{R}$ defined by $\varphi(x, \alpha) = f(x) - \frac{1}{2}\alpha x^2$. If $\varphi(x, \alpha)$ is a convex function on J for some $\alpha = \alpha^*$, then $\varphi(x, \alpha)$ is called a **convexification** of f and α^* a **convexifier** on J. A function f is **convexifiable** if it has a convexification. It is noted in [244, Corollary 2.9] that if the continuous differentiable function f has Lipschitz derivative (i. e., $|f'(x) - f'(y)| \le L|x - y|$ for any $x, y \in J$ and some constant L), then $\alpha = -L$ is a convexifier of f.

The following fact concerning convexifiable function plays an important role in our discussion (see [244, Corollary 2.8]):

If f is twice continuous differentiable, then $\alpha = \min_{t \in J} f''(t)$ is a convexifier of f. **(P)**

After the above preparation, we are ready to prove the analogue of (5.1.1) for non-convex functions.

Theorem 5.1.2 (Operator Jensen inequality for nonconvex function [178]). *Let f be a continuous convexifiable function on the interval J and α a convexifier of f. Then*

$$f(\langle Ax, x\rangle) \le \langle f(A)x, x\rangle - \frac{1}{2}\alpha(\langle A^2x, x\rangle - \langle Ax, x\rangle^2), \tag{5.1.2}$$

for every self-adjoint operator A with $Sp(A) \subseteq J$ and every unit vector $x \in \mathcal{H}$.

Proof. Let $g_\alpha : J \to \mathbb{R}$ with $g_\alpha(x) = \varphi(x, \alpha)$. According to the assumption, $g_\alpha(x)$ is convex. Therefore,

$$g_\alpha(\langle Ax, x\rangle) \le \langle g_\alpha(A)x, x\rangle,$$

for every unit vector $x \in \mathcal{H}$. This expression is equivalent to the desired inequality (5.1.2). □

Theorem 5.1.2 can be regarded as a *generalization* of Theorem 5.1.1 in the sense that Theorem 5.1.2 never assumes the convexity of the function f. A few remarks concerning Theorem 5.1.2 are in order.

Remark 5.1.1.

(i) Using the fact that for a convex function f, one can choose the convexifier $\alpha = 0$; one recovers the inequality (5.1.1). That is, (5.1.2) is an extension of (5.1.1).

(ii) For continuous differentiable function f with Lipschitz derivative and Lipschitz constant L, we have

$$f(\langle Ax,x\rangle) \le \langle f(A)x,x\rangle + \frac{1}{2}L(\langle A^2x,x\rangle - \langle Ax,x\rangle^2).$$

An important special case of Theorem 5.1.2, which refines inequality (5.1.1) can be explicitly stated using the property (**P**).

Remark 5.1.2. Let $f : J \to \mathbb{R}$ be a twice continuous differentiable strictly convex function and $\alpha = \min_{t\in J} f''(t)$. Then

$$f(\langle Ax,x\rangle) \le \langle f(A)x,x\rangle - \frac{1}{2}\alpha(\langle A^2x,x\rangle - \langle Ax,x\rangle^2) \le \langle f(A)x,x\rangle, \tag{5.1.3}$$

for every positive operator A with $Sp(A) \subseteq J$ and every unit vector $x \in \mathcal{H}$.

Corollary 5.1.1. *Let f be a continuous convexifiable function on the interval J and α a convexifier. Let $A_1,\dots,A_n$ be self-adjoint operators on $\mathcal{H}$ with $Sp(A_i) \subseteq J$ for $i = 1,\dots,n$ and $x_1,\dots,x_n \in \mathcal{H}$ be such that $\sum_{i=1}^n \|x_i\|^2 = 1$. Then we have*

$$f\left(\sum_{i=1}^n \langle A_ix_i,x_i\rangle\right) \le \sum_{i=1}^n \langle f(A_i)x_i,x_i\rangle - \frac{1}{2}\alpha\left(\sum_{i=1}^n \langle A_i^2x_i,x_i\rangle - \left(\sum_{i=1}^n \langle A_ix_i,x_i\rangle\right)^2\right). \tag{5.1.4}$$

Proof. In fact, $\mathrm{x} := (x_1,\dots,x_n)^T$ is a unit vector in the Hilbert space $\mathcal{H}^n$. If we introduce the *diagonal* operator on $\mathcal{H}^n$,

$$\Lambda := \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_n \end{pmatrix},$$

then, obviously, $Sp(\Lambda) \subseteq J$, $\|\mathrm{x}\| = 1$, $\langle f(\Lambda)\mathrm{x},\mathrm{x}\rangle = \sum_{i=1}^n \langle f(A_i)x_i,x_i\rangle$, $\langle \Lambda\mathrm{x},\mathrm{x}\rangle = \sum_{i=1}^n \langle A_ix_i,x_i\rangle$, $\langle \Lambda^2\mathrm{x},\mathrm{x}\rangle = \sum_{i=1}^n \langle A_i^2x_i,x_i\rangle$. Hence, to complete the proof, it is enough to apply Theorem 5.1.2 for Λ and x. □

Corollary 5.1.1 leads us to the following result.

Corollary 5.1.2. *Let f be a continuous convexifiable function on the interval J and α a convexifier. Let $A_1,\dots,A_n$ be self-adjoint operators on $\mathcal{H}$ with $Sp(A_i) \subseteq J$ for $i = 1,\dots,n$ and let $p_1,\dots,p_n$ be positive numbers such that $\sum_{i=1}^n p_i = 1$. Then we have*

$$f\left(\sum_{i=1}^n \langle p_iA_ix,x\rangle\right) \le \sum_{i=1}^n \langle p_if(A_i)x,x\rangle - \frac{1}{2}\alpha\left(\sum_{i=1}^n \langle p_iA_i^2x,x\rangle - \left(\sum_{i=1}^n \langle p_iA_ix,x\rangle\right)^2\right), \tag{5.1.5}$$

for every unit vector $x \in \mathcal{H}$.

Proof. Suppose that $x \in \mathcal{H}$ is a unit vector. Putting $x_i = \sqrt{p_i}x \in \mathcal{H}$ so that $\sum_{i=1}^{n} \|x_i\|^2 = 1$ and applying Corollary 5.1.1 we obtain the desired result (5.1.5). □

The clear advantage of our approach over the Jensen operator inequality is shown in the following example. Before proceeding, we recall the following multiple operator version of Jensen's inequality [3, Corollary 1]: Let $f : [m, M] \subseteq \mathbb{R} \to \mathbb{R}$ be a convex function and A_i be self-adjoint operators with $Sp(A_i) \subseteq [m, M]$, $i = 1, \dots, n$ for some scalars $m < M$. If $p_i \geq 0$, $i = 1, \dots, n$ with $\sum_{i=1}^{n} p_i = 1$, then

$$f\left(\sum_{i=1}^{n} \langle p_i A_i x, x\rangle\right) \leq \sum_{i=1}^{n} \langle p_i f(A_i)x, x\rangle, \tag{5.1.6}$$

for every unit vector $x \in \mathcal{H}$.

Example 5.1.1. Let $f(t) = \sin t (0 \leq t \leq 2\pi)$, $\alpha = \min_{0 \leq t \leq 2\pi} f''(t) = -1$, $n = 2$, $p_1 = p$, $p_2 = 1 - p$, $\mathcal{H} = \mathbb{R}^2$, $A_1 = \left(\begin{smallmatrix} 2\pi & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $A_2 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 2\pi \end{smallmatrix}\right)$ and $x = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$. Then we have $\varphi''(t, -1) = 1 - \sin t \geq 0$ for $\varphi(t, -1) = f(t) + \frac{1}{2}t^2$. After simple calculations with (5.1.5), we infer that

$$\sin(2\pi(1-p)) \leq 2\pi^2 p(1-p), \quad 0 \leq p \leq 1 \tag{5.1.7}$$

and (5.1.6) implies

$$\sin(2\pi(1-p)) \leq 0, \quad 0 \leq p \leq 1. \tag{5.1.8}$$

It is not a surprise matter that the inequality (5.1.8) does not hold when $\frac{1}{2} \leq p \leq 1$. That is, (5.1.6) is not applicable here. However, we may emphasize that the new upper bound in (5.1.7) still holds.

It follows from Corollary 5.1.2 that

$$f\left(\sum_{i=1}^{n} p_i t_i\right) \leq \sum_{i=1}^{n} p_i f(t_i) - \frac{1}{2}\alpha\left(\sum_{i=1}^{n} p_i t_i^2 - \left(\sum_{i=1}^{n} p_i t_i\right)^2\right), \tag{5.1.9}$$

where $t_i \in J$ and $\sum_{i=1}^{n} p_i = 1$. For the case $n = 2$, the inequality (5.1.9) reduces to

$$f((1-v)t_1 + vt_2) \leq (1-v)f(t_1) + vf(t_2) - \frac{v(1-v)}{2}\alpha(t_1 - t_2)^2, \tag{5.1.10}$$

where $0 \leq v \leq 1$. In particular,

$$f\left(\frac{t_1 + t_2}{2}\right) \leq \frac{f(t_1) + f(t_2)}{2} - \frac{1}{8}\alpha(t_1 - t_2)^2. \tag{5.1.11}$$

It is notable that Theorem 5.1.2 is equivalent to the inequality (5.1.9). The following provides a refinement of the arithmetic–geometric mean inequality.

Proposition 5.1.1. *For each $a, b > 0$ and $v \in [0,1]$, we have*

$$\sqrt{ab} \le H_v(a,b) - \frac{d}{8}\left((1-2v)\left(\log\frac{a}{b}\right)\right)^2 \le \frac{a+b}{2} - \frac{d}{8}\left(\log\frac{a}{b}\right)^2 \le \frac{a+b}{2}, \tag{5.1.12}$$

where $d = \min\{a, b\}$ and $H_v(a,b) = \frac{a^{1-v}b^v + b^{1-v}a^v}{2}$ is the **Heinz mean.**

Proof. Assume that f is a twice differentiable convex function such that $\alpha \le f''$ where $\alpha \in \mathbb{R}$. Under these conditions, it follows that

$$\begin{aligned}
f\left(\frac{a+b}{2}\right) &= f\left(\frac{(1-v)a + vb + (1-v)b + va}{2}\right) \\
&\le \frac{f((1-v)a + vb) + f((1-v)b + va)}{2} - \frac{1}{8}\alpha((a-b)(1-2v))^2 \quad \text{(by (5.1.11))} \\
&\le \frac{f(a) + f(b)}{2} - \frac{1}{8}\alpha(a-b)^2 \quad \text{(by (5.1.10))} \\
&\le \frac{f(a) + f(b)}{2},
\end{aligned}$$

for $\alpha \ge 0$. Now taking $f(t) = e^t$ with $t \in J = [a, b]$ in the above inequalities, we deduce the desired inequality (5.1.12). □

Remark 5.1.3. As R. Bhatia pointed out in [21], the Heinz means interpolate between the geometric mean and the arithmetic mean, that is,

$$\sqrt{ab} \le H_v(a,b) \le \frac{a+b}{2}. \tag{5.1.13}$$

Of course, the first inequality in (5.1.12) yields an improvement of (5.1.13). The inequalities in (5.1.12) also sharpens up the following inequality which is due to S. S. Dragomir (see [41, Remark 1]):

$$\frac{d}{8}\left(\log\frac{a}{b}\right)^2 \le \frac{a+b}{2} - \sqrt{ab}.$$

Studying about the arithmetic–geometric mean inequality, we cannot avoid mentioning its cousin, the Young inequality. The following **Zuo–Liao inequality** provides a ratio refinement and reverse of the Young inequality:

$$K^r(h) \le \frac{(1-v)a + vb}{a^{1-v}b^v} \le K^R(h), \tag{5.1.14}$$

where $v \in [0,1]$, $r = \min\{v, 1-v\}$, $R = \max\{v, 1-v\}$ and $h = \frac{b}{a}$. The first one was proved by H. Zuo et al. [248, Corollary 3], while the second one was given by W. Liao et al. [140, Corollary 2.2].

Our aim in the following is to establish a refinement for the inequalities in (5.1.14). The crucial role for our purposes will play the following facts:

If f is a convex function on the fixed closed interval J, then we have

$$np_{\min}\left\{\sum_{i=1}^{n}\frac{1}{n}f(x_i)-f\left(\sum_{i=1}^{n}\frac{1}{n}x_i\right)\right\}\le\sum_{i=1}^{n}p_if(x_i)-f\left(\sum_{i=1}^{n}p_ix_i\right),\tag{5.1.15}$$

$$\sum_{i=1}^{n}p_if(x_i)-f\left(\sum_{i=1}^{n}p_ix_i\right)\le np_{\max}\left\{\sum_{i=1}^{n}\frac{1}{n}f(x_i)-f\left(\sum_{i=1}^{n}\frac{1}{n}x_i\right)\right\},\tag{5.1.16}$$

where $p_1,\dots,p_n\ge 0$ with $\sum_{i=1}^{n}p_i=1$, $p_{\min}=\min\{p_1,\dots,p_n\}$, $p_{\max}=\max\{p_1,\dots,p_n\}$. Notice that the first inequality goes back to J. Pečarić et al. [165, Theorem 1], while the second one was obtained by F.-C. Mitroi in [166, Corollary 3.1]. We here find that the inequality (5.1.15) is a generalization of Proposition 2.2.1 in Section 2.2.

We state the theorem. In order to simplify the notation, we put $a\sharp_v b=a^{1-v}b^v$ and $a\nabla_v b=(1-v)a+vb$.

Theorem 5.1.3 ([178]). *Let $a,b>0$ and $v\in[0,1]$. Then we have*

$$\begin{aligned}K^r(h)\exp\left(\left(\frac{v(1-v)}{2}-\frac{r}{4}\right)\left(\frac{a-b}{D}\right)^2\right)\\ \le\frac{a\nabla_v b}{a\sharp_v b}\le K^R(h)\exp\left(\left(\frac{v(1-v)}{2}-\frac{R}{4}\right)\left(\frac{a-b}{D}\right)^2\right),\end{aligned}\tag{5.1.17}$$

where $r=\min\{v,1-v\}$, $R=\max\{v,1-v\}$, $D=\max\{a,b\}$ and $h=\frac{b}{a}$.

Proof. Employing the inequality (5.1.15) for the twice differentiable convex function f with $\alpha\le f''$, we have

$$\begin{aligned}&np_{\min}\left\{\frac{1}{n}\sum_{i=1}^{n}f(x_i)-f\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)\right\}-\sum_{i=1}^{n}p_if(x_i)+f\left(\sum_{i=1}^{n}p_ix_i\right)\\ &\le\frac{\alpha}{2}\left\{np_{\min}\left[\frac{1}{n}\sum_{i=1}^{n}x_i^2-\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)^2\right]-\left(\sum_{i=1}^{n}p_ix_i^2-\left(\sum_{i=1}^{n}p_ix_i\right)^2\right)\right\}.\end{aligned}$$

Here, we set $n=2$, $x_1=a$, $x_2=b$, $p_1=1-v$, $p_2=v$, $v=r$ and $f(x)=-\log x$ with $J=[a,b]$, then $\alpha=\min_{x\in J}f''(x)=\frac{1}{D^2}$. Thus we deduce the first inequality in (5.1.17). The second inequality in (5.1.17) is also obtained similarly by using the inequality (5.1.16). □

Remark 5.1.4.

(i) Since $\frac{v(1-v)}{2}-\frac{r}{4}\ge 0$ for $v\in[0,1]$, we have $\exp((\frac{v(1-v)}{2}-\frac{r}{4})(\frac{a-b}{D})^2)\ge 1$. Therefore, the first inequality in (5.1.17) provides an improvement for the first inequality in (5.1.14).

(ii) Since $\frac{v(1-v)}{2}-\frac{R}{4}\le 0$ for $v\in[0,1]$, we get $\exp((\frac{v(1-v)}{2}-\frac{R}{4})(\frac{a-b}{D})^2)\le 1$. Therefore, the second inequality in (5.1.17) provides an improvement for the second inequality in (5.1.14).

Proposition 5.1.2. *Under the same assumptions in Theorem* 5.1.3, *we have*

$$\frac{(h+1)^2}{4h} \geq \exp\left(\frac{1}{4}\left(\frac{a-b}{D}\right)^2\right).$$

Proof. We prove the case $a \leq b$, then $h \geq 1$. We set $f_1(h) = 2\log(h+1) - \log h - 2\log 2 - \frac{1}{4}\frac{(h-1)^2}{h^2}$. It is quite easy to see that $f_1'(h) = \frac{(2h+1)(h-1)^2}{2h^3(h+1)} \geq 0$, so that $f_1(h) \geq f_1(1) = 0$. For the case $a \geq b$, (then $0 < h \leq 1$), we also set $f_2(h) = 2\log(h+1) - \log h - 2\log 2 - \frac{1}{4}(h-1)^2$. By direct calculation $f_2'(h) = -\frac{(h-1)^2(h+2)}{2h(h+1)} \leq 0$, so that $f_2(h) \geq f_2(1) = 0$. Thus the statement follows. □

Remark 5.1.5. S. S. Dragomir obtained a refinement and reverse of Young inequality in [41, Theorem 3] as

$$\exp\left(\frac{v(1-v)}{2}\left(\frac{a-b}{D}\right)^2\right) \leq \frac{a\nabla_v b}{a\sharp_v b} \leq \exp\left(\frac{v(1-v)}{2}\left(\frac{a-b}{d}\right)^2\right), \tag{5.1.18}$$

where $d = \min\{a, b\}$. From the following facts (i) and (ii), we claim that our inequalities are nontrivial results:

(i) From Proposition 5.1.2, our lower bound in (5.1.17) is tighter than the one in (5.1.18).

(ii) Numerical computations show that there is no ordering between the right-hand side in (5.1.17) and the one in the second inequality of (5.1.18) shown in [41, Theorem 3]. See [178] for the example.

We give a further remark in relation to comparisons with other inequalities.

Remark 5.1.6. The following refined Young inequality and its reverse are known

$$K^{r'}(\sqrt{t}, 2)t^v + r(1-\sqrt{t})^2 \leq (1-v) + vt \leq K^{R'}(\sqrt{t}, 2)t^v + r(1-\sqrt{t})^2, \tag{5.1.19}$$

where $t > 0$, $v \in [0, 1]$, $r = \min\{v, 1-v\}$, $r' = \min\{2r, 1-2r\}$ and $R' = \max\{2r, 1-2r\}$. The first and second inequality were given in [235, Lemma 2.1] and in [140, Theorem 2.1], respectively. Numerical computations show that there is no ordering between our inequalities (5.1.17) and the above ones. See [178] for the details.

Obviously, in the inequality (5.1.14), we cannot replace $K^r(h)$ by $K^R(h)$, or vice versa. In this regard, we have the following theorem. The proof is almost same as that of Theorem 5.1.3 (it is enough to use the convexity of the function $g_\beta(x) = \frac{\beta}{2}x^2 - f(x)$ where $\beta = \max_{x \in J} f''(x)$).

Theorem 5.1.4 ([178]). *Let all the assumptions of Theorem* 5.1.3 *hold with* $d := \min\{a, b\}$. *Then*

$$\begin{aligned} K^R(h)\exp\left(\left(\frac{v(1-v)}{2} - \frac{R}{4}\right)\left(\frac{a-b}{d}\right)^2\right) &\leq \frac{a\nabla_v b}{a\sharp_v b} \\ &\leq K^r(h)\exp\left(\left(\frac{v(1-v)}{2} - \frac{r}{4}\right)\left(\frac{a-b}{d}\right)^2\right). \end{aligned}$$

We end this section by presenting the operator inequalities based on Theorems 5.1.3 and 5.1.4. See [178] for the proof.

Corollary 5.1.3. *Let A, B be two positive invertible operators and positive real numbers m, m', M, M' that satisfy one of the following conditions:*
(i) $0 < m'I \le A \le mI < MI \le B \le M'I$.
(ii) $0 < m'I \le B \le mI < MI \le A \le M'I$.

Then

$$\begin{aligned} &K^r(h)\exp\left(\left(\frac{\nu(1-\nu)}{2}-\frac{r}{4}\right)\left(\frac{1-h}{h}\right)^2\right)A\sharp_\nu B \\ &\quad \le A\nabla_\nu B \le K^R(h')\exp\left(\left(\frac{\nu(1-\nu)}{2}-\frac{R}{4}\right)\left(\frac{1-h'}{h'}\right)^2\right)A\sharp_\nu B \end{aligned} \tag{5.1.20}$$

and

$$\begin{aligned} &K^R(h)\exp\left(\left(\frac{\nu(1-\nu)}{2}-\frac{R}{4}\right)\left(\frac{1-h'}{h'}\right)^2\right)A\sharp_\nu B \\ &\quad \le A\nabla_\nu B \le K^r(h')\exp\left(\left(\frac{\nu(1-\nu)}{2}-\frac{r}{4}\right)\left(\frac{1-h}{h}\right)^2\right)A\sharp_\nu B, \end{aligned}$$

where $r = \min\{\nu, 1-\nu\}$, $R = \max\{\nu, 1-\nu\}$, $h = \frac{M}{m}$ and $h' = \frac{M'}{m'}$.

5.2 Choi–Davis–Jensen inequality without convexity

C. Davis [38] and M. D. Choi [34] showed that if $\Phi : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ is a normalized positive linear map and if f is an operator convex function on an interval J, then the so-called **Choi–Davis–Jensen inequality** (in short C-D-J inequality)

$$f(\Phi(A)) \le \Phi(f(A)) \tag{5.2.1}$$

holds for every self-adjoint operator A on $\mathcal{H}$ whose spectrum is contained in J.

The inequality (5.2.1) can break down when the operator convexity is dropped. For instance, taking

$$A = \begin{pmatrix} 4 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}, \quad \Phi\left((a_{ij})_{1\le i,j\le 3}\right) = (a_{ij})_{1\le i,j\le 2} \quad \text{and} \quad f(t) = t^4.$$

By a simple computation, we have

$$\begin{pmatrix} 325 & 132 \\ 132 & 61 \end{pmatrix} = \Phi^4(A) \nleq \Phi(A^4) = \begin{pmatrix} 374 & 105 \\ 105 & 70 \end{pmatrix}.$$

This example shows that the inequality (5.2.1) will be false if we replace the operator convex function by a usual convex function. In [160, Theorem 1], J. Mićić et al. pointed out that the inequality (5.2.1) holds true for real valued continuous convex functions with conditions on the bounds of the operators.

The purpose of this section is to obtain the C-D-J inequality for nonconvex functions. Applying our main results, we give new inequalities improving previous known results such as the Kantorovich inequality, and bounds for relative operator entropies and quantum mechanical entropies. As for applications to entropy theory, see Chapter 7.

Theorem 5.2.1 ([159]). *Let $f : J \to \mathbb{R}$ be continuous twice differentiable function such that $\alpha \le f'' \le \beta$ where $\alpha, \beta \in \mathbb{R}$ and let $\Phi : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ be a normalized positive linear map. Then we have*

$$f(\Phi(A)) \le \Phi(f(A)) + \frac{\beta-\alpha}{2}\{(M+m)\Phi(A) - Mm\} + \frac{1}{2}(\alpha\Phi(A)^2 - \beta\Phi(A^2)), \tag{5.2.2}$$

and

$$\Phi(f(A)) \le f(\Phi(A)) + \frac{\beta-\alpha}{2}\{(M+m)\Phi(A) - Mm\} + \frac{1}{2}(\alpha\Phi(A^2) - \beta\Phi(A)^2), \tag{5.2.3}$$

for any self-adjoint operator A on $\mathcal{H}$ with the spectrum $Sp(A) \subseteq [m, M] \subset J$.

In order to prove Theorem 5.2.1, we need the following lemma.

Lemma 5.2.1. *Let $f : J \to \mathbb{R}$ be continuous twice differentiable function such that $\alpha \le f'' \le \beta$, where $\alpha, \beta \in \mathbb{R}$, and let $\Phi : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ be a normalized positive linear map. If A is a self-adjoint operator on $\mathcal{H}$ with $Sp(A) \subseteq [m, M] \subset J$ for some $m < M$, then we have*

$$\Phi(f(A)) \le L(\Phi(A)) - \frac{\alpha}{2}\{(M+m)\Phi(A) - Mm - \Phi(A^2)\}, \tag{5.2.4}$$

$$\Phi(f(A)) \ge L(\Phi(A)) - \frac{\beta}{2}\{(M+m)\Phi(A) - Mm - \Phi(A^2)\}, \tag{5.2.5}$$

$$f(\Phi(A)) \le L(\Phi(A)) - \frac{\alpha}{2}\{(M+m)\Phi(A) - Mm - \Phi(A)^2\}, \tag{5.2.6}$$

$$f(\Phi(A)) \ge L(\Phi(A)) - \frac{\beta}{2}\{(M+m)\Phi(A) - Mm - \Phi(A)^2\}, \tag{5.2.7}$$

where

$$L(t) := \frac{M-t}{M-m}f(m) + \frac{t-m}{M-m}f(M), \tag{5.2.8}$$

is the line that passes through the points $(m, f(m))$ and $(M, f(M))$.

Proof. Since $\alpha \le f'' \le \beta$, then the function $g_\alpha(x) := f(x) - \frac{\alpha}{2}x^2$ is convex. So,

$$g_\alpha((1-v)a + vb) \le (1-v)g_\alpha(a) + vg_\alpha(b),$$

holds for any $0 \le v \le 1$ and $a, b \in J$. It follows that

$$f\big((1-v)a + vb\big) \le (1-v)f(a) + vf(b) - \frac{\alpha}{2}\, v(1-v)(a-b)^2.$$

Since any $t \in [m, M]$ can be written in the form $t = \frac{M-t}{M-m}m + \frac{t-m}{M-m}M$, and putting $v = \frac{t-m}{M-m}$, $a = m$ and $b = M$ in the above inequality we have

$$f(t) \le L(t) - \frac{\alpha}{2}\big((M+m)t - mM - t^2\big). \tag{5.2.9}$$

Now, by using the standard functional calculus of a self-adjoint operator A to (5.2.9) and next applying a normalized positive linear map Φ we obtain

$$\Phi\big(f(A)\big) \le \Phi\big(L(A)\big) - \frac{\alpha}{2}\{(M+m)\Phi(A) - Mm - \Phi\big(A^2\big)\},$$

which gives the desired inequality (5.2.4).

By applying (5.2.9) on $\Phi(A)$, we obtain (5.2.6). Using the same technique as above for a convex function $g_\beta(t) := \frac{\beta}{2}x^2 - f(x)$, we obtain

$$L(t) - \frac{\beta}{2}\big((M+m)t - mM - t^2\big) \le f(t),$$

which gives (5.2.5) and (5.2.7). □

From Lemma 5.2.1, we can prove Theorem 5.2.1.

Proof of Theorem 5.2.1. Let $m < M$. We obtain (5.2.2) after combining (5.2.6) with (5.2.5) and we obtain (5.2.3) after combining (5.2.4) with (5.2.7). □

Remark 5.2.1. The inequality (5.2.3) is a reverse of C-D-J inequality $f(\Phi(A)) \le \Phi(f(A))$ for a nonconvex function. The second term in (5.2.3) is always nonnegative, while the sign of third term in (5.2.3) is not determined.

Example 5.2.1. To illustrate Theorem 5.2.1 works properly, let $\Phi(A) = \langle Ax, x\rangle$,($x$ is a unit vector), where $x = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$, $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 3 & 1 \\ -1 & 1 & 2 \end{pmatrix}$ and $f(t) = t^3$. Of course, we can choose $m = 0.25$ and $M = 3.8$. So after some calculations, we see that

$$\begin{aligned} 8 &= f\big(\Phi(A)\big) \\ &\lneq \Phi\big(f(A)\big) + \frac{\beta-\alpha}{2}\{(M+m)\Phi(A) - Mm\} + \frac{1}{2}\big(\alpha\Phi(A)^2 - \beta\Phi\big(A^2\big)\big) \approx 27.14, \end{aligned}$$

and

$$\begin{aligned} 24 &= \Phi\big(f(A)\big) \\ &\lneq f\big(\Phi(A)\big) + \frac{\beta-\alpha}{2}\{(M+m)\Phi(A) - Mm\} + \frac{1}{2}\big(\alpha\Phi\big(A^2\big) - \beta\Phi(A)^2\big) \approx 43.54. \end{aligned}$$

Theorem 5.2.2 ([159]). *Let A be a self-adjoint operator with $Sp(A) \subseteq [m,M] \subset J$ for some $m < M$. If $f : [m,M] \to (0,\infty)$ is a continuous twice differentiable function such that $\alpha \leq f''$, where $\alpha \in \mathbb{R}$, and if $\Phi : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ is a normalized positive linear map, then we have*

$$\frac{1}{K(m,M,f)}\left\{\Phi(f(A)) + \frac{\alpha}{2}[(M+m)\Phi(A) - Mm - \Phi(A^2)]\right\}$$
$$\leq f(\Phi(A)) \leq K(m,M,f)\Phi(f(A)) - \frac{\alpha}{2}[(M+m)\Phi(A) - Mm - \Phi(A)^2], \tag{5.2.10}$$

where

$$K(m,M,f) = \max\left\{\frac{L(t)}{f(t)} : t \in [m,M]\right\} \tag{5.2.11}$$

and $L(t)$ is defined in (5.2.8).

Proof. By using (5.2.6), we have (see [164, Corollary 4.12])

$$f(\Phi(A)) \leq L(\Phi(A)) - \frac{\alpha}{2}[(M+m)\Phi(A) - Mm - \Phi(A)^2]$$
$$\leq K(m,M,f)\Phi(f(A)) - \frac{\alpha}{2}[(M+m)\Phi(A) - Mm - \Phi(A)^2],$$

which gives the right-hand side in the inequality of (5.2.10). Also, by using (5.2.4) and given that $0 < m \leq \Phi(A) < M$, we obtain

$$\Phi(f(A)) \leq K(m,M,f)f(\Phi(A)) - \frac{\alpha}{2}[(M+m)\Phi(A) - Mm - \Phi(A^2)].$$

Since $K(m,M,f) > 0$, it follows:

$$f(\Phi(A)) \geq \frac{1}{K(m,M,f)}\left\{\Phi(f(A)) + \frac{\alpha}{2}[(M+m)\Phi(A) - Mm - \Phi(A^2)]\right\},$$

which is the left-hand side in the inequality of (5.2.10). □

Remark 5.2.2. Let A and Φ be as in Theorem 5.2.2. If $f : [m,M] \to (0,\infty)$ is a continuous twice differentiable function such that $f'' \leq \beta$, where $\beta \in \mathbb{R}$, then by using (5.2.5) and (5.2.7), we can obtain the following result:

$$k(m,M,f)\Phi(f(A)) - \frac{\beta}{2}[(M+m)\Phi(A) - Mm - \Phi(A)^2]$$
$$\leq f(\Phi(A)) \leq \frac{1}{k(m,M,f)}\left\{\Phi(f(A)) + \frac{\beta}{2}[(M+m)\Phi(A) - Mm - \Phi(A^2)]\right\},$$

where $k(m,M,f) = \min\{\frac{L(t)}{f(t)} : t \in [m,M]\}$ and $L(t)$ is defined by (5.2.8).

In the next corollary, we give a refinement of reverse of C-D-J inequality (see, e. g., [164])

$$\frac{1}{K(m,M,f)}\Phi(f(A)) \le f(\Phi(A)) \le K(m,M,f)\Phi(f(A)),$$

for every strictly convex function f on $[m,M]$, where $K(m,M,f) > 1$ is defined by (5.2.11).

Corollary 5.2.1. *Under the assumptions of Theorem* 5.2.2, *for a strictly convex function* f *on* $[m,M]$. *Then we have*

$$\begin{aligned}\frac{1}{K(m,M,f)}\Phi(f(A)) &\le \frac{1}{K(m,M,f)}\left\{\Phi(f(A)) + \frac{\alpha}{2}[(M+m)\Phi(A) - Mm - \Phi(A^2)]\right\}\\ &\le f(\Phi(A)) \le K(m,M,f)\Phi(f(A)) - \frac{\alpha}{2}[(M+m)\Phi(A) - Mm - \Phi(A)^2]\\ &\le K(m,M,f)\Phi(f(A)),\end{aligned}$$

where $K(m,M,f) > 1$ *is defined by* (5.2.11).

Proof. Since f is strictly convex, then $0 < \alpha \le f''$ on $[m,M]$. Given that $(M+m)\Phi(A) - Mm - \Phi(A^2) \ge 0$ is valid, the desired result follows by applying Theorem 5.2.2. □

5.3 Operator Jensen–Mercer inequality

Our motivation for this section arose from the paper by A. McD. Mercer [157], which is connected with a remarkable variant of the inequality (5.0.2). His result states the following.

Theorem 5.3.1 ([157, Theorem 1.2]). *If* f *is a convex function on* $[m,M]$, *then we have*

$$f\left(M + m - \sum_{i=1}^{n} w_i x_i\right) \le f(M) + f(m) - \sum_{i=1}^{n} w_i f(x_i), \tag{5.3.1}$$

for all $x_i \in [m,M]$ *and all* $w_i \in [0,1]$ $(i = 1,\dots,n)$ *with* $\sum_{i=1}^{n} w_i = 1$.

A similar arguing with application of functional calculus gives an **operator Jensen–Mercer inequality** without operator convexity assumptions. More precisely, the following theorem is proved in [154].

Theorem 5.3.2 ([154, Theorem 1]). *Let* $A_1,\dots,A_n \in \mathbb{B}(\mathcal{H})$ *be self-adjoint operators with spectrum in* $[m,M]$ *and let* $\Phi_1,\dots,\Phi_n : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ *be positive linear maps with* $\sum_{i=1}^{n}\Phi_i(\mathbf{1}_{\mathcal{H}}) = \mathbf{1}_{\mathcal{K}}$. *If* $f : [m,M] \subseteq \mathbb{R} \to \mathbb{R}$ *is convex function, then we have*

$$f\left((M+m)1_K - \sum_{i=1}^{n}\Phi_i(A_i)\right) \le (f(M) + f(m))1_K - \sum_{i=1}^{n}\Phi_i(f(A_i)). \tag{5.3.2}$$

Moreover, in the same paper [154], the following series of inequalities was proved:

$$
\begin{aligned}
f\left((M+m)\mathbf{1}_{\mathcal{K}} - \sum_{i=1}^{n}\Phi_i(A_i)\right) &\le (f(M)+f(m))\mathbf{1}_{\mathcal{K}} + \frac{\sum_{i=1}^{n}\Phi_i(A_i) - M\mathbf{1}_{\mathcal{K}}}{M-m} f(m) + \frac{m\mathbf{1}_{\mathcal{K}} - \sum_{i=1}^{n}\Phi_i(A_i)}{M-m} f(M) \\
&\le (f(M)+f(m))\mathbf{1}_{\mathcal{K}} - \sum_{i=1}^{n}\Phi_i(f(A_i)). \qquad (5.3.3)
\end{aligned}
$$

In the following, we aim to improve inequality (5.3.1). The following lemma is well known in [157, Lemma 1.3], but we prove it for the reader's convenience.

Lemma 5.3.1. *Let f be a convex function on $[m,M]$, then*

$$f(M+m-a_i) \le f(M)+f(m)-f(a_i), \quad (m \le a_i \le M, i = 1,\dots,n).$$

Proof. If $f : [m,M] \to \mathbb{R}$ is a convex function, then for any $x, y \in [m,M]$ and $t \in [0,1]$, we have

$$f(tx+(1-t)y) \le tf(x)+(1-t)f(y). \tag{5.3.4}$$

It can be verified that if $m \le a_i \le M$ $(i = 1,\dots,n)$, then $\frac{M-a_i}{M-m}, \frac{a_i-m}{M-m} \le 1$ and $\frac{M-a_i}{M-m} + \frac{a_i-m}{M-m} = 1$. Thanks to (5.3.4), we have

$$f(a_i) \le \frac{M-a_i}{M-m} f(m) + \frac{a_i-m}{M-m} f(M). \tag{5.3.5}$$

One the other hand, $m \le a_i \le M$ $(i = 1,\dots,n)$ implies $m \le M+m-a_i \le M$ $(i = 1,\dots,n)$. Thus, from (5.3.5) we infer

$$f(M+m-a_i) \le \frac{a_i-m}{M-m} f(m) + \frac{M-a_i}{M-m} f(M). \tag{5.3.6}$$

Summing up (5.3.5) and (5.3.6), we get the desired result. □

Based on this, our first result can be stated as follows.

Theorem 5.3.3 ([173, Theorem 2.1]). *Let f be a convex function on $[m,M]$ and $t \in [0,1]$. Then*

$$
\begin{aligned}
f(M+m-\overline{a}) &\le \sum_{i=1}^{n} w_i f(M+m-((1-t)\overline{a}+ta_i)) \\
&\le f(M)+f(m) - \sum_{i=1}^{n} w_i f(a_i) \qquad (5.3.7)
\end{aligned}
$$

for all $a_i \in [m, M]$ *and* $w_i \in [0,1]$ $(i = 1, \ldots, n)$ *with* $\sum_{i=1}^{n} w_i = 1$, *where* $\bar{a} := \sum_{i=1}^{n} w_i a_i$. *Moreover, the function* $F : [0,1] \to \mathbb{R}$ *defined by*

$$F(t) = \sum_{i=1}^{n} w_i f\big(M + m - \big((1-t)\overline{a} + ta_i\big)\big)$$

is monotonically nondecreasing and convex on $[0,1]$.

Proof. First, we have

$$\begin{aligned}\sum_{i=1}^{n} w_i f\big(M + m - \big((1-t)\overline{a} + ta_i\big)\big) &\geq f\left(\sum_{i=1}^{n} w_i\big(M + m - \big((1-t)\overline{a} + ta_i\big)\big)\right) \\ &= f(M + m - \overline{a}).\end{aligned} \tag{5.3.8}$$

On the other hand,

$$\begin{aligned}&\sum_{i=1}^{n} w_i f\big(M + m - \big((1-t)\overline{a} + ta_i\big)\big) \\ &\quad = \sum_{i=1}^{n} w_i f\big((1-t)(M + m - \overline{a}) + t(M + m - a_i)\big) \\ &\quad \leq \sum_{i=1}^{n} w_i\big((1-t)f(M + m - \overline{a}) + tf(M + m - a_i)\big) \\ &\quad \leq \sum_{i=1}^{n} w_i\left((1-t)\left(f(M) + f(m) - \sum_{j=1}^{n} w_j f(a_j)\right) + t\big(f(M) + f(m) - f(a_i)\big)\right) \\ &\quad = f(M) + f(m) - \sum_{i=1}^{n} w_i f(a_i).\end{aligned}$$

For the convexity of F, we have

$$\begin{aligned}F\left(\frac{t+s}{2}\right) &= \sum_{i=1}^{n} w_i f\left(M + m - \left(\left(1 - \frac{t+s}{2}\right)\overline{a} + \frac{t+s}{2}a_i\right)\right) \\ &= \sum_{i=1}^{n} w_i f\left(M + m - \left(\frac{(1-t)\overline{a} + ta_i + (1-s)\overline{a} + sa_i}{2}\right)\right) \\ &= \sum_{i=1}^{n} w_i f\left(\frac{M + m - ((1-t)\overline{a} + ta_i) + M + m - ((1-s)\overline{a} + sa_i)}{2}\right) \\ &\leq \frac{1}{2}\left[\sum_{i=1}^{n} w_i f\big(M + m - \big((1-t)\overline{a} + ta_i\big)\big) + \sum_{i=1}^{n} w_i f\big(M + m - \big((1-s)\overline{a} + sa_i\big)\big)\right] \\ &= \frac{F(t) + F(s)}{2}.\end{aligned}$$

Now, if $0 < s < t < 1$, then $s = \frac{t-s}{t} \cdot 0 + \frac{s}{t} \cdot t$, and hence the convexity of F implies

$$\begin{aligned} F(s) &= F\left(\frac{t-s}{t} \cdot 0 + \frac{s}{t} \cdot t\right) \\ &\le \frac{t-s}{t} F(0) + \frac{s}{t} F(t) \\ &\le \frac{t-s}{t} F(t) + \frac{s}{t} F(t) \\ &= F(t). \end{aligned}$$

We remark that the second inequality in the above follows from (5.3.8) and the fact

$$F(0) = \sum_{i=1}^{n} w_i f(M + m - \overline{a}) = f(M + m - \overline{a}).$$

Therefore, F is monotonically nondecreasing on $[0, 1]$. □

Corollary 5.3.1. *Let all the assumptions of Theorem* 5.3.3 *hold, then*

$$f(M + m - \overline{a}) \le \sum_{i=1}^{n} \frac{w_i}{\overline{a} - a_i} \int_{M+m-\overline{a}}^{M+m-a_i} f(t)\, dt \le f(M) + f(m) - \sum_{i=1}^{n} w_i f(a_i).$$

Proof. Integrating the inequality (5.3.7) over $t \in [0, 1]$, we get Corollary 5.3.1. Here, we used the fact

$$\begin{aligned} \int_0^1 f\big(M + m - ((1-t)\overline{a} + t a_i)\big)\, dt &= \int_0^1 f\big((1-t)(M + m - \overline{a}) + t(M + m - a_i)\big)\, dt \\ &= \int_0^1 f\big(t(M + m - \overline{a}) + (1-t)(M + m - a_i)\big)\, dt \\ &= \frac{1}{\overline{a} - a_i} \int_{M+m-\overline{a}}^{M+m-a_i} f(t)\, dt. \end{aligned}$$ □

Remark 5.3.1. Put $n = 2$, $w_1 = w_2 = 1/2$, $a_1 = a$, and $a_2 = b$ in Corollary 5.3.1, then

$$\begin{aligned} f\left(M + m - \frac{a+b}{2}\right) &\le \frac{1}{b-a} \int_a^b f(M + m - u)\, du \\ &\le f(M) + f(m) - \frac{f(a) + f(b)}{2}. \end{aligned}$$

We give a more precise estimate in the next theorem.

Theorem 5.3.4 ([173, Theorem 2.2]). *Let f be a convex function on $[m, M]$. Then*

$$\begin{aligned} f(M+m-\overline{a}) &\le \sum_{i=1}^{n} w_i f\left(M+m-\frac{\overline{a}+a_i}{2}\right) \\ &\le \sum_{i=1}^{n} \frac{w_i}{\overline{a}-a_i} \int_{M+m-\overline{a}}^{M+m-a_i} f(t)\,dt \\ &\le f(M)+f(m)-\sum_{i=1}^{n} w_i f(a_i) \end{aligned} \tag{5.3.9}$$

for all $a_i \in [m, M]$ and $w_i \in [0,1]$ $(i = 1, \dots, n)$ with $\sum_{i=1}^{n} w_i = 1$, where $\bar{a} := \sum_{i=1}^{n} w_i a_i$.

Proof. If $f : [m, M] \to \mathbb{R}$ is a convex function, then we have for any $a, b \in [m, M]$,

$$\begin{aligned} f\left(\frac{a+b}{2}\right) &= f\left(\frac{ta+(1-t)b+tb+(1-t)a}{2}\right) \\ &\le \frac{f(ta+(1-t)b)+f(tb+(1-t)a)}{2} \\ &\le \frac{f(a)+f(b)}{2}. \end{aligned}$$

Replacing a and b by $M+m-a$ and $M+m-b$, respectively, we get

$$\begin{aligned} f\left(M+m-\frac{a+b}{2}\right) &\le \frac{f(M+m-(ta+(1-t)b))+f(M+m-(tb+(1-t)a))}{2} \\ &\le \frac{f(M+m-a)+f(M+m-b)}{2}. \end{aligned}$$

Integrating the inequalities over $t \in [0,1]$, and using the fact

$$\int_0^1 f(tx+(1-t)y)\,dt = \int_0^1 f(ty+(1-t)x)\,dt,$$

we infer that

$$\begin{aligned} f\left(M+m-\frac{a+b}{2}\right) &\le \int_0^1 f\big(M+m-(ta+(1-t)b)\big)\,dt \\ &\le \frac{f(M+m-a)+f(M+m-b)}{2}. \end{aligned}$$

Since $a_i, \bar{a} \in [m, M]$, we can write

$$\begin{aligned} f\left(M+m-\frac{\overline{a}+a_i}{2}\right) &\le \frac{1}{\overline{a}-a_i} \int_{M+m-\overline{a}}^{M+m-a_i} f(t)\,dt \\ &\le \frac{f(M+m-\overline{a})+f(M+m-a_i)}{2}, \end{aligned}$$

due to

$$\int_0^1 f(M+m-(t\overline{a}+(1-t)a_i))\,dt = \frac{1}{\overline{a}-a_i}\int_{M+m-\overline{a}}^{M+m-a_i} f(t)\,dt.$$

Multiplying by $w_i > 0$ ($i = 1,\ldots,n$) and summing over i from 1 to n, we may deduce

$$\begin{aligned}\sum_{i=1}^{n} w_i f\left(M+m-\frac{\overline{a}+a_i}{2}\right) &\le \sum_{i=1}^{n} \frac{w_i}{\overline{a}-a_i}\int_{M+m-\overline{a}}^{M+m-a_i} f(t)\,dt\\ &\le \frac{f(M+m-\overline{a})+\sum_{i=1}^{n} w_i f(M+m-a_i)}{2}.\end{aligned} \tag{5.3.10}$$

On the other hand, by (5.0.2)

$$\begin{aligned}f(M+m-\bar{a}) &= f\left(\sum_{i=1}^{n} w_i\left(M+m-\frac{\bar{a}+a_i}{2}\right)\right)\\ &\le \sum_{i=1}^{n} w_i f\left(M+m-\frac{\bar{a}+a_i}{2}\right)\end{aligned} \tag{5.3.11}$$

and by Lemma 5.3.1 and (5.3.1)

$$\begin{aligned}&\frac{f(M+m-\bar{a})+\sum_{i=1}^{n} w_i f(M+m-a_i)}{2}\\ &\quad\le \frac{f(M)+f(m)-\sum_{j=1}^{n} w_j f(a_j)+f(M)+f(m)-\sum_{i=1}^{n} w_i f(a_i)}{2}\\ &\quad= f(M)+f(m)-\sum_{i=1}^{n} w_i f(a_i).\end{aligned} \tag{5.3.12}$$

Combining (5.3.10), (5.3.11) and (5.3.12), we get (5.3.9). □

Corollary 5.3.2. *Let $a_i \in [m,M]$ and $w_i \in [0,1]$ ($i = 1,\ldots,n$) with $\sum_{i=1}^{n} w_i = 1$. Then*

$$\begin{aligned}\frac{Mm}{\prod_{i=1}^{n} a_i^{w_i}} &\le \exp\left[\sum_{i=1}^{n}\frac{w_i}{\overline{a}-a_i}\int_{M+m-\overline{a}}^{M+m-a_i}\log t\,dt\right]\\ &\le \prod_{i=1}^{n}\left(M+m-\frac{\overline{a}+a_i}{2}\right)^{w_i}\\ &\le M+m-\sum_{i=1}^{n} w_i a_i.\end{aligned}$$

Proof. Put $f(t) = -\log t$, $(0 < t \le 1)$ in Theorem 5.3.4. □

Remark 5.3.2. If we set $n = 2$, $a_1 = m$, $a_2 = M$ and $w_1 = w_2 = 1/2$ in Corollary 5.3.2, then we have

$$\sqrt{Mm} \le \frac{M^{\frac{M}{M-m}}}{em^{\frac{m}{M-m}}} \le \frac{1}{4}\sqrt{(M+3m)(m+3M)} \le \frac{1}{2}(M+m).$$

In the following theorem, we give an **extension of operator Jensen–Mercer inequality** given in Theorem 5.3.2.

Theorem 5.3.5 ([174]). *Let $A_1,\dots,A_n \in \mathbb{B}(\mathcal{H})$ be self-adjoint operators with spectrum in $[m,M]$ and let $\Phi_1,\dots,\Phi_n : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ be positive linear maps with $\sum_{i=1}^{n}\Phi_i(\mathbf{1}_{\mathcal{H}}) = \mathbf{1}_{\mathcal{K}}$. If $f : [m,M] \subseteq \mathbb{R} \to \mathbb{R}$ is a continuous twice differentiable function such that $\alpha \le f'' \le \beta$ with $\alpha,\beta \in \mathbb{R}$, then we have*

$$\begin{aligned}
&(f(M)+f(m))\mathbf{1}_{\mathcal{K}} - \sum_{i=1}^{n}\Phi_i(f(A_i))\\
&\quad - \beta\left\{(M+m)\sum_{i=1}^{n}\Phi_i(A_i) - Mm\mathbf{1}_{\mathcal{K}} - \frac{1}{2}\left(\left(\sum_{i=1}^{n}\Phi_i(A_i)\right)^2 + \sum_{i=1}^{n}\Phi_i(A_i^2)\right)\right\}\\
&\le f\left((M+m)\mathbf{1}_{\mathcal{K}} - \sum_{i=1}^{n}\Phi_i(A_i)\right) && (5.3.13)\\
&\le (f(M)+f(m))\mathbf{1}_{\mathcal{K}} - \sum_{i=1}^{n}\Phi_i(f(A_i))\\
&\quad - \alpha\left\{(M+m)\sum_{i=1}^{n}\Phi_i(A_i) - Mm\mathbf{1}_{\mathcal{K}} - \frac{1}{2}\left(\left(\sum_{i=1}^{n}\Phi_i(A_i)\right)^2 + \sum_{i=1}^{n}\Phi_i(A_i^2)\right)\right\}. && (5.3.14)
\end{aligned}$$

Proof. The idea of the proof is the following. We give the sketch of the proof. It is well known that for any convex function f and $m \le t \le M$ we have

$$f(t) = f\left(\frac{M-t}{M-m}m + \frac{t-m}{M-m}M\right) \le L(t), \tag{5.3.15}$$

where

$$L(t) = \frac{M-t}{M-m}f(m) + \frac{t-m}{M-m}f(M). \tag{5.3.16}$$

According to the assumption, the function $g_\alpha(t) = f(t) - \frac{\alpha}{2}t^2$ $(m \le t \le M)$ is convex. On account of (5.3.15), we have

$$f(t) \le L(t) - \frac{\alpha}{2}\{(M+m)t - Mm - t^2\}. \tag{5.3.17}$$

Since $m \le M+m-t \le M$, we can replace t with $M+m-t$, which gives us

$$f(M+m-t) \le L_0(t) - \frac{\alpha}{2}\{(M+m)t - Mm - t^2\},$$

where $L_0(t) = L(M + m - t) = f(M) + f(m) - L(t)$. Using functional calculus for the operator $m\mathbf{1}_{\mathcal{K}} \le \sum_{i=1}^{n} \Phi_i(A_i) \le M\mathbf{1}_{\mathcal{K}}$, we infer that

$$\begin{aligned} & f\left((M+m)\mathbf{1}_{\mathcal{K}} - \sum_{i=1}^{n} \Phi_i(A_i) \right) \\ & \quad \le L_0\left(\sum_{i=1}^{n} \Phi_i(A_i) \right) - \frac{\alpha}{2} \left\{ (M+m) \sum_{i=1}^{n} \Phi_i(A_i) - Mm\mathbf{1}_{\mathcal{K}} - \left(\sum_{i=1}^{n} \Phi_i(A_i) \right)^2 \right\}. \end{aligned} \tag{5.3.18}$$

On the other hand, by applying functional calculus for the operator $m\mathbf{1}_{\mathcal{H}} \le A_i \le M\mathbf{1}_{\mathcal{H}}$ in (5.3.17), we get

$$f(A_i) \le L(A_i) - \frac{\alpha}{2}\{(M+m)A_i - Mm\mathbf{1}_{\mathcal{H}} - A_i^2\}.$$

Applying positive linear maps Φ_i and summing, we have

$$\begin{aligned} & \sum_{i=1}^{n} \Phi_i(f(A_i)) \\ & \quad \le L\left(\sum_{i=1}^{n} \Phi_i(A_i) \right) - \frac{\alpha}{2} \left\{ (M+m) \sum_{i=1}^{n} \Phi_i(A_i) - Mm\mathbf{1}_{\mathcal{K}} - \sum_{i=1}^{n} \Phi_i(A_i^2) \right\}. \end{aligned} \tag{5.3.19}$$

Combining the two inequalities (5.3.18) and (5.3.19), we get (5.3.14). The inequality (5.3.13) follows similarly by taking into account that

$$L(t) - \frac{\beta}{2}\{(M+m)t - Mm - t^2\} \le f(t), \quad m \le t \le M.$$

The details are left to the reader. Hence, we have the conclusion. □

This expression has the advantage of using twice differentiable functions instead of convex functions used in Theorem 5.3.2. Here, we give an example to clarify the situation in Theorem 5.3.5.

Example 5.3.1. Taking $f(t) = \sin t (0 \le t \le 2\pi)$, $A = \begin{pmatrix} \frac{\pi}{4} & 0 \\ 0 & \frac{\pi}{2} \end{pmatrix}$ and $\Phi(A) = \frac{1}{2}\operatorname{Tr}[A]$. After simple computations (by putting $m = \frac{\pi}{4}$ and $M = \frac{\pi}{2}$), we get

$$0.9238 \approx f\big((M+m) - \Phi(A)\big) \nleq f(M) + f(m) - \Phi\big(f(A)\big) \approx 0.8535.$$

This example shows that (5.3.2) may fail without the convexity assumption. On the other hand,

$$\begin{aligned} 0.9238 & \approx f\big((M+m) - \Phi(A)\big) \\ & \lneqq f(M) + f(m) - \Phi\big(f(A)\big) - \alpha\left\{ (M+m)\Phi(A) - Mm - \frac{1}{2}\{\Phi(A)^2 + \Phi(A^2)\} \right\} \\ & \approx 0.9306, \end{aligned}$$

namely our approach can fill the gap.

It is important that, under convexity assumption, a strong result related to Theorem 5.3.2 hold.

Remark 5.3.3. It is instructive to observe that

$$(M+m)\sum_{i=1}^{n}\Phi_i(A_i) - Mm\mathbf{1}_{\mathcal{K}} - \frac{1}{2}\left(\left(\sum_{i=1}^{n}\Phi_i(A_i)\right)^2 + \sum_{i=1}^{n}\Phi_i(A_i^2)\right) \geq 0.$$

One the other hand, if f is convex then $\alpha \geq 0$. This shows that (5.3.14) can provide a much stronger bound than (5.3.2). (Of course, the inequality (5.3.18) is also sharper than (5.3.3).)

In the sequel, we will briefly review some known properties related to log-convex functions. A positive function defined on an interval (or, more generally, on a convex subset of some vector space) is called **log-convex** if $\log f(x)$ is a convex function of x. We observe that such functions satisfy the elementary inequality

$$f\big((1-v)a+vb\big) \leq f^{1-v}(a)f^{v}(b), \quad v \in [0,1]$$

for any $a, b \in J$ and a nonnegative function f. The function f is called **log-concave** if the inequality above works in the reversed way (i. e., when f^{-1} is log-convex). Because of the arithmetic–geometric mean inequality, we also have

$$f\big((1-v)a+vb\big) \leq f^{1-v}(a)f^{v}(b) \leq (1-v)f(a)+vf(b), \tag{5.3.20}$$

which says that any log-convex function is a convex function. This is of interest to us because (5.3.20) can be written as

$$f(t) \leq f^{\frac{M-t}{M-m}}(m)f^{\frac{t-m}{M-m}}(M) \leq L(t), \quad m \leq t \leq M \tag{5.3.21}$$

where $L(t)$ is as in (5.3.17). With the inequality (5.3.21), we can present the following result, which can be regarded as an extension of Theorem 5.3.2 to log-convex functions. The proof is left to the reader as an exercise.

Theorem 5.3.6 ([174]). *Let all the assumptions of Theorem* 5.3.5 *hold except that* $f : [m,M] \to (0,\infty)$ *is a* log-*convex. Then we have*

$$\begin{aligned} f\left((M+m)\mathbf{1}_{\mathcal{K}} - \sum_{i=1}^{n}\Phi_i(A_i)\right) &\leq [f(m)]^{\frac{\sum_{i=1}^{n}\Phi_i(A_i)-m\mathbf{1}_{\mathcal{K}}}{M-m}}[f(M)]^{\frac{M\mathbf{1}_{\mathcal{K}}-\sum_{i=1}^{n}\Phi_i(A_i)}{M-m}} \\ &\leq \big(f(M)+f(m)\big)\mathbf{1}_{\mathcal{K}} - \sum_{i=1}^{n}\Phi_i\big(f(A_i)\big). \end{aligned} \tag{5.3.22}$$

In the rest of this section, we give some applications to the results given in this section. Assume that:

(i) $\mathbf{A} = (A_1, \dots, A_n)$, where $A_i \in \mathbb{B}(\mathcal{H})$ are positive invertible operators with $Sp(A_i) \subseteq [m, M]$ for some scalars $0 < m < M$.
(ii) $\mathbf{\Phi} = (\Phi_1, \dots, \Phi_n)$, where $\Phi_i : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ are positive linear maps.

Here, $C([m, M])$ is the set of all real valued continuous functions on an interval $[m, M]$. In [154], the following expression is defined, where the authors called the **operator quasi-arithmetic mean of Mercer's type**:

$$\widetilde{M}_\varphi(\mathbf{A}, \mathbf{\Phi}) = \varphi^{-1}\left((\varphi(M) + \varphi(m))\mathbf{1}_\mathcal{K} - \sum_{i=1}^n \Phi_i(\varphi(A_i)) \right).$$

Theorem 5.3.7 ([174]). *Let $\varphi, \psi \in C([m, M])$ be two strictly monotone functions.*
(i) *If either $\psi o \varphi^{-1}$ is convex and ψ^{-1} is operator monotone increasing, or $\psi o \varphi^{-1}$ is concave and ψ^{-1} is operator monotone decreasing, then*

$$\widetilde{M}_\varphi(\mathbf{A}, \mathbf{\Phi}) \le \widetilde{M}_\psi(\mathbf{A}, \mathbf{\Phi}). \tag{5.3.23}$$

(ii) *If either $\psi o \varphi^{-1}$ is concave and ψ^{-1} is operator monotone increasing, or $\psi o \varphi^{-1}$ is convex and ψ^{-1} is operator monotone decreasing, then the inequality in* (5.3.23) *is reversed.*

These interesting inequalities were firstly discovered by A. Matković et al. [154, Theorem 4]. By virtue of Theorem 5.3.5, we have the following result.

Theorem 5.3.8 ([174]). *Let $\varphi, \psi \in C([m, M])$ be two strictly monotone functions and $\psi o \varphi^{-1}$ is twice differentiable function.*
(i) *If $\alpha \le (\psi o \varphi^{-1})''$ with $\alpha \in \mathbb{R}$ and ψ^{-1} is operator monotone, then*

$$\widetilde{M}_\varphi(\mathbf{A}, \mathbf{\Phi}) \le \psi^{-1}\{\psi(\widetilde{M}_\psi(\mathbf{A}, \mathbf{\Phi})) - \alpha \diamond (m, M, \varphi, \mathbf{A}, \mathbf{\Phi})\}, \tag{5.3.24}$$

where

$$\begin{aligned}\diamond(m, M, \varphi, \mathbf{A}, \mathbf{\Phi}) &= (\varphi(M) + \varphi(m)) \sum_{i=1}^n \Phi_i(\varphi(A_i)) - \varphi(M)\varphi(m)\mathbf{1}_\mathcal{K} \\ &\quad - \frac{1}{2}\left(\left(\sum_{i=1}^n \Phi_i(\varphi(A_i)) \right)^2 + \sum_{i=1}^n \Phi_i(\varphi(A_i)^2) \right).\end{aligned}$$

(ii) *If $(\psi o \varphi^{-1})'' \le \beta$ with $\beta \in \mathbb{R}$ and ψ^{-1} is operator monotone, then the reverse inequality is valid in* (5.3.24) *with β instead of α.*

Proof. Make the substitution $f = \psi o \varphi^{-1}$ in (5.3.14) and replace A_i, m and M with $\varphi(A_i)$, $\varphi(m)$ and $\varphi(M)$, respectively, we get

$$\psi(\widetilde{M}_\varphi(\mathbf{A}, \mathbf{\Phi})) \le \psi(\widetilde{M}_\psi(\mathbf{A}, \mathbf{\Phi})) - \alpha \diamond (m, M, \varphi, \mathbf{A}, \mathbf{\Phi}).$$

Since ψ^{-1} is operator monotone, we can get the conclusion. The other case follows in a similar manner from (5.3.13). □

In the same spirit, we infer from Theorem 5.3.6 the following result:

Theorem 5.3.9 ([174]). *Let $\varphi, \psi \in C([m, M])$ be two strictly monotone functions. If $\psi o \varphi^{-1}$ is* log*-convex function and ψ^{-1} is operator increasing, then*

$$\widetilde{M}_{\varphi}(\mathbf{A}, \mathbf{\Phi}) \le \psi^{-1}\{[\psi(m)]^{\frac{\sum_{i=1}^{n}\Phi_i(\varphi(A_i))-\varphi(m)\mathbf{1}_{\mathcal{K}}}{\varphi(M)-\varphi(m)}}[\psi(M)]^{\frac{\varphi(M)\mathbf{1}_{\mathcal{K}}-\sum_{i=1}^{n}\Phi_i(\varphi(A_i))}{\varphi(M)-\varphi(m)}}\} \le \widetilde{M}_{\psi}(\mathbf{A}, \mathbf{\Phi}).$$

Remark 5.3.4. By choosing adequate functions φ and ψ, and appropriate substitutions, we can obtain some improvements concerning operator power mean of Mercer's type. We leave the details of this idea to the interested reader, as it is just an application of main results in this section.

In the end of this section, we show the example such that there is no relationship between Theorems 5.3.8 and 5.3.9. Here, we restrict ourselves to the power function $f(t) = t^p$ with $p < 0$.

Example 5.3.2. It is sufficient to compare (5.3.17) and the first inequality of (5.3.21). We take $m = 1$, $M = 3$. Setting

$$g(t) = \frac{M-t}{M-m}m^p + \frac{t-m}{M-m}M^p - \frac{p(p-1)M^{p-2}}{2}\{(M+m)t - Mm - t^2\} - \left(m^{\frac{M-t}{M-m}}M^{\frac{t-m}{M-m}}\right)^p.$$

A simple calculation shows $g(2) \approx -0.0052909$ when $p = -0.2$, while $g(2) \approx 0.0522794$ when $p = -1$. We thus conclude that there is no ordering the right-hand side of (5.3.17) and the first inequality of (5.3.21).

5.4 Inequalities for operator concave function

During the past decades several formulations, extensions or refinements of the Kantorovich inequality [153] in various settings have been introduced by many mathematicians. See [143, 164, 182, 187] and references therein.

Let $A \in \mathbb{B}(\mathcal{H})$ be a positive operator such that $mI \le A \le MI$ for some scalars $0 < m < M$ and Φ be a normalized positive linear map on $\mathbb{B}(\mathcal{H})$, then

$$\Phi(A^{-1}) \sharp \Phi(A) \le \frac{M+m}{2\sqrt{Mm}}. \tag{5.4.1}$$

In addition,

$$\Phi(A) \sharp \Phi(B) \le \frac{M+m}{2\sqrt{Mm}}\Phi(A \sharp B), \tag{5.4.2}$$

whenever $m^2 A \le B \le M^2 A$ and $0 < m < M$. The first inequality goes back to R. Nakamoto and M. Nakamura in [189], the second is more general and has been proved in [138] by E.-Y. Lee as a matrix version.

In the below, we first *extend* (5.4.2), then as an application, we obtain a *generalization* of (5.4.1). In addition, we use elementary operations and give some inequalities related to the Bellman type. We prove the following *new* result, from which (5.4.2) directly follows.

Theorem 5.4.1 ([222]). *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two strictly positive operators such that $m_1^2 I \le A \le M_1^2 I$, $m_2^2 I \le B \le M_2^2 I$ for some positive scalars $m_1 < M_1$, $m_2 < M_2$ and let Φ be a normalized positive linear map on $\mathbb{B}(\mathcal{H})$. If $f : [0,\infty) \to [0,\infty)$ is an operator monotone, then we have*

$$f\left(\left(\frac{M+m}{2}\right)\Phi(A\sharp B)\right) \ge f\left(\frac{Mm\Phi(A)+\Phi(B)}{2}\right) \ge \frac{f(Mm\Phi(A))+f(\Phi(B))}{2}$$
$$\ge f(Mm\Phi(A))\sharp f(\Phi(B)),$$

where $m = \frac{m_2}{M_1}$ and $M = \frac{M_2}{m_1}$.

Proof. According to the assumptions, we have $mI \le (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}} \le MI$, it follows that $(M+m)(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}} \ge MmI + A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$. The above inequality then implies $(\frac{M+m}{2})A\sharp B \ge \frac{MmA+B}{2}$. Using the hypotheses made about Φ, we have $(\frac{M+m}{2})\Phi(A\sharp B) \ge \frac{Mm\Phi(A)+\Phi(B)}{2}$. Thus we have

$$\begin{aligned} f\left(\left(\frac{M+m}{2}\right)\Phi(A\sharp B)\right) &\ge f\left(\frac{Mm\Phi(A)+\Phi(B)}{2}\right) \quad \text{(since } f \text{ is operator monotone)} \\ &\ge \frac{f(Mm\Phi(A))+f(\Phi(B))}{2} \quad \text{(by [8, Corollary 1.12])} \\ &\ge f(Mm\Phi(A))\sharp f(\Phi(B)) \quad \text{(by AM-GM inequality)}, \end{aligned}$$

which is the statement of the theorem. □

We complement Theorem 5.4.1 by proving the following.

Theorem 5.4.2 ([222]). *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two strictly positive operators such that $m_1^2 I \le A \le M_1^2 I$, $m_2^2 I \le B \le M_2^2 I$ for some scalars $m_1 < M_1$, $m_2 < M_2$, and let Φ be a normalized positive linear map on $\mathbb{B}(\mathcal{H})$. If $g : [0,\infty) \to [0,\infty)$ is operator monotone decreasing, then*

$$\begin{aligned} g\left(\left(\frac{M+m}{2}\right)\Phi(A\sharp B)\right) &\le g\left(\frac{Mm\Phi(A)+\Phi(B)}{2}\right) \\ &\le \left\{\frac{g(Mm\Phi(A))^{-1}+g(\Phi(B))^{-1}}{2}\right\}^{-1} \le g(Mm\Phi(A))\sharp g(\Phi(B)), \end{aligned}$$

where $m = \frac{m_2}{M_1}$ and $M = \frac{M_2}{m_1}$.

Proof. Since g is operator monotone decreasing on $(0,\infty)$, so $\frac{1}{g}$ is operator monotone on $(0,\infty)$. Now by applying Theorem 5.4.1 for $f = \frac{1}{g}$, we have

$$\begin{aligned} g\Big(\Big(\frac{M+m}{2}\Big)\Phi(A\sharp B)\Big)^{-1} &\geq g\Big(\frac{Mm\Phi(A)+\Phi(B)}{2}\Big)^{-1} \\ &\geq \frac{g(Mm\Phi(A))^{-1}+g(\Phi(B))^{-1}}{2} \geq g\big(Mm\Phi(A)\big)^{-1}\sharp g\big(\Phi(B)\big)^{-1}. \end{aligned}$$

Taking the inverse, we get

$$\begin{aligned} g\Big(\Big(\frac{M+m}{2}\Big)\Phi(A\sharp B)\Big) &\leq g\Big(\frac{Mm\Phi(A)+\Phi(B)}{2}\Big) \leq \Big\{\frac{g(Mm\Phi(A))^{-1}+g(\Phi(B))^{-1}}{2}\Big\}^{-1} \\ &\leq \big\{g\big(Mm\Phi(A)\big)^{-1}\sharp g\big(\Phi(B)\big)^{-1}\big\}^{-1} = g\big(Mm\Phi(A)\big)\sharp g\big(\Phi(B)\big), \end{aligned}$$

proving the main assertion of the theorem. □

As a byproduct of Theorems 5.4.1 and 5.4.2, we have the following result.

Corollary 5.4.1. *Under the assumptions of Theorem* 5.4.1, *we have the following:*
(i) *If* $0 \leq r \leq 1$, *then*

$$\Big(\frac{M+m}{2\sqrt{Mm}}\Big)^r\Phi^r(A\sharp B) \geq \Big(\frac{Mm\Phi(A)+\Phi(B)}{2\sqrt{Mm}}\Big)^r \geq \frac{(Mm)^r\Phi^r(A)+\Phi^r(B)}{2(Mm)^{\frac{r}{2}}} \geq \Phi^r(A)\sharp\Phi^r(B).$$

(ii) *If* $-1 \leq r \leq 0$, *then*

$$\begin{aligned} \Big(\frac{M+m}{2\sqrt{Mm}}\Big)^r\Phi^r(A\sharp B) &\leq \Big(\frac{Mm\Phi(A)+\Phi(B)}{2\sqrt{Mm}}\Big)^r \leq \frac{1}{(Mm)^{\frac{r}{2}}}\Big\{\frac{(Mm)^{-r}\Phi^{-r}(A)+\Phi^{-r}(B)}{2}\Big\}^{-1} \\ &\leq \Phi^r(A)\sharp\Phi^r(B). \end{aligned}$$

Our next result is a straightforward application of Theorems 5.4.1 and 5.4.2.

Corollary 5.4.2. *Let* $A \in \mathbb{B}(\mathcal{H})$ *be a positive operator such that* $mI \leq A \leq MI$ *for some scalars* $0 < m < M$ *and* Φ *be a normalized positive linear map on* $\mathbb{B}(\mathcal{H})$.
(i) *If* $f : [0,\infty) \to [0,\infty)$ *is operator monotone, then*

$$\begin{aligned} f\Big(\frac{M+m}{2Mm}\Big) &\geq f\Big(\frac{\frac{1}{Mm}\Phi(A)+\Phi(A^{-1})}{2}\Big) \geq \frac{f(\frac{1}{Mm}\Phi(A))+f(\Phi(A^{-1}))}{2} \\ &\geq f\Big(\frac{1}{Mm}\Phi(A)\Big)\sharp f(\Phi(A^{-1})). \end{aligned}$$

(ii) *If* $g : [0,\infty) \to [0,\infty)$ *is operator monotone decreasing, then*

$$\begin{aligned} g\Big(\frac{M+m}{2Mm}\Big) &\leq g\Big(\frac{\frac{1}{Mm}\Phi(A)+\Phi(A^{-1})}{2}\Big) \leq \Big\{\frac{g(\frac{1}{Mm}\Phi(A))^{-1}+g(\Phi(A^{-1}))^{-1}}{2}\Big\}^{-1} \\ &\leq g\Big(\frac{1}{Mm}\Phi(A)\Big)\sharp g(\Phi(A^{-1})). \end{aligned}$$

In the same vein as in Corollary 5.4.1, we have the following consequences.

Corollary 5.4.3. *Under the assumptions of Corollary 5.4.2, we have the following:*
(i) *If* $0 \le r \le 1$, *then*

$$\left(\frac{M+m}{2\sqrt{Mm}}\right)^r \ge \left(\frac{\frac{1}{\sqrt{Mm}}\Phi(A) + \sqrt{Mm}\Phi(A^{-1})}{2}\right)^r \ge \frac{\frac{1}{(Mm)^{\frac{r}{2}}}\Phi^r(A) + (Mm)^{\frac{r}{2}}\Phi^r(A^{-1})}{2}$$
$$\ge \Phi^r(A)\sharp\Phi^r(A^{-1}).$$

For the special case of $r = 1$, *we have*

$$\frac{M+m}{2\sqrt{Mm}} \ge \frac{\frac{1}{\sqrt{Mm}}\Phi(A) + \sqrt{Mm}\Phi(A^{-1})}{2} \ge \Phi(A)\sharp\Phi(A^{-1}).$$

(ii) *If* $-1 \le r \le 0$, *then*

$$\left(\frac{M+m}{2\sqrt{Mm}}\right)^r \le \left(\frac{\frac{1}{\sqrt{Mm}}\Phi(A) + \sqrt{Mm}\Phi(A^{-1})}{2}\right)^r \le \left\{\frac{(Mm)^r\Phi^{-r}(A) + \Phi^{-r}(A^{-1})}{2(Mm)^{\frac{r}{2}}}\right\}^{-1}$$
$$\le \Phi^r(A)\sharp\Phi^r(A^{-1}).$$

Let $A, B \in \mathbb{B}(\mathcal{H})$ be two strictly positive operators and Φ be a normalized positive linear map on $\mathbb{B}(\mathcal{H})$. If $f : [0,\infty) \to [0,\infty)$ is an operator concave, then for any $v \in [0,1]$, the following inequality obtained in [185, Theorem 2.1]:

$$\Phi(f(A))\nabla_v\Phi(f(B)) \le f(\Phi(A\nabla_v B)). \tag{5.4.3}$$

In the same paper [185], as an **operator Bellman inequality** [17], M. S. Moslehian et al. showed that

$$\Phi((I-A)^r\nabla_v(I-B)^r) \le \Phi^r(I - A\nabla_v B), \tag{5.4.4}$$

where A, B are two **norm-contractive** (in the sense that $\|A\|, \|B\| \le 1$) and $r, v \in [0,1]$.
Under the convexity assumption on f, (5.4.4) can be reversed.

Theorem 5.4.3 ([222]). *Let* $A, B \in \mathbb{B}(\mathcal{H})$ *be two norm contractive and* Φ *be a normalized positive linear map on* $\mathbb{B}(\mathcal{H})$. *Then we have*

$$\Phi^r(I - A\nabla_v B) \le \Phi((I-A)^r\nabla_v(I-B)^r), \tag{5.4.5}$$

for any $v \in [0,1]$ *and* $r \in [-1,0] \cup [1,2]$.

Proof. If f is operator convex, we have

$$\begin{aligned} f(\Phi(A\nabla_v B)) &\le \Phi(f(A\nabla_v B)) \quad \text{(by Choi–Davis–Jensen inequality)} \\ &\le \Phi(f(A)\nabla_v f(B)) \quad \text{(by operator convexity of } f). \end{aligned}$$

The function $f(t) = t^r$ is operator convex on $(0,\infty)$ for $r \in [-1,0] \cup [1,2]$. It can be verified that $f(t) = (1-t)^r$ is operator convex on $(0,1)$ for $r \in [-1,0] \cup [1,2]$. This implies the desired result (5.4.5). □

However, we are looking for something stronger than (5.4.5). The principal object of this section is to prove the following.

Theorem 5.4.4 ([222]). *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two contraction operators and Φ be a normalized positive linear map on $\mathbb{B}(\mathcal{H})$. Then we have*

$$\Phi^r(I - A\nabla_v B) \le \Phi^r(I-A)\sharp_v \Phi^r(I-B) \le \Phi\big((I-A)^r \sharp_v (I-B)^r\big) \le \Phi\big((I-A)^r \nabla_v (I-B)^r\big),$$

where $v \in [0,1]$ and $r \in [-1,0]$.

The proof is given at the end of this section. The following lemma will play an important role in our proof.

Lemma 5.4.1. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two strictly positive operators and Φ be a normalized positive linear map on $\mathbb{B}(\mathcal{H})$. If $f : [0,\infty) \to [0,\infty)$ is operator monotone decreasing, then for any $v \in [0,1]$*

$$f\big(\Phi(A\nabla_v B)\big) \le f\big(\Phi(A)\big)\sharp_v f\big(\Phi(B)\big) \le \Phi\big(f(A)\big)\nabla_v \Phi\big(f(B)\big), \tag{5.4.6}$$

and

$$f\big(\Phi(A\nabla_v B)\big) \le \Phi\big(f(A)\sharp_v f(B)\big) \le \Phi\big(f(A)\big)\nabla_v \Phi\big(f(B)\big). \tag{5.4.7}$$

More precisely,

$$f\big(\Phi(A\nabla_v B)\big) \le f\big(\Phi(A)\big)\sharp_v f\big(\Phi(B)\big) \le \Phi\big(f(A)\sharp_v f(B)\big) \le \Phi\big(f(A)\big)\nabla_v \Phi\big(f(B)\big). \tag{5.4.8}$$

Proof. As T. Ando and F. Hiai mentioned in [8, (2.16)], the function f is operator monotone decreasing if and only if

$$f(A\nabla_v B) \le f(A)\sharp_v f(B). \tag{5.4.9}$$

We emphasize here that if f satisfies (5.4.9), then it is operator convex (this class of functions is called **operator log-convex**). It is easily verified that if $Sp(A), Sp(B) \subseteq J$, then $Sp(\Phi(A)), Sp(\Phi(B)) \subseteq J$. So we can replace A, B by $\Phi(A), \Phi(B)$ in (5.4.9), respectively. Therefore, we can write

$$\begin{aligned} f\big(\Phi(A\nabla_v B)\big) &\le f\big(\Phi(A)\big)\sharp_v f\big(\Phi(B)\big) \\ &\le \Phi\big(f(A)\big)\sharp_v \Phi\big(f(B)\big) \quad \text{(by C-D-J inequality and monotonicity property of mean)} \\ &\le \Phi\big(f(A)\nabla_v f(B)\big) \quad \text{(by AM-GM inequality).} \end{aligned}$$

This completes the proof of the inequality (5.4.6). To prove the inequality (5.4.7), note that if $Sp(A), Sp(B) \subseteq J$, then $Sp(A\nabla_v B) \subseteq J$. By computation,

$$\begin{aligned} f(\Phi(A\nabla_v B)) &\leq \Phi(f(A\nabla_v B)) \quad \text{(by Choi–Davis–Jensen inequality)} \\ &\leq \Phi(f(A)\sharp_v f(B)) \quad \text{(by (5.4.9))} \\ &\leq \Phi(f(A))\sharp_v \Phi(f(B)) \quad \text{(by Ando's inequality [6, Theorem 3])} \\ &\leq \Phi(f(A)\nabla_v f(B)) \quad \text{(by AM-GM inequality),} \end{aligned}$$

proving the inequality (5.4.7). We know that if $g > 0$ is operator monotone on $(0, \infty)$, then g is operator concave. As before, it can be shown that

$$g(\Phi(A))\sharp_v g(\Phi(B)) \geq \Phi(g(A))\sharp_v \Phi(g(B)) \geq \Phi(g(A)\sharp_v g(B)).$$

Taking the inverse, we get

$$g(\Phi(A))^{-1}\sharp_v g(\Phi(B))^{-1} \leq \Phi(g(A)\sharp_v g(B))^{-1} \leq \Phi(g(A)^{-1}\sharp_v g(B)^{-1}).$$

If g is operator monotone, then $f = \frac{1}{g}$ is operator monotone decreasing, we conclude

$$f(\Phi(A))\sharp_v f(\Phi(B)) \leq \Phi(f(A)\sharp_v f(B)).$$

This proves (5.4.8). □

We finally present the proof of Theorem 5.4.4.

Proof of Theorem 5.4.4. It is well known that the function $f(t) = t^r$ on $(0, \infty)$ is operator monotone decreasing for $r \in [-1, 0]$. It implies that the function $f(t) = (1-t)^r$ on $(0, 1)$ is also operator monotone decreasing. By applying Lemma 5.4.1, we get the desired result. □

5.5 Bounds on operator Jensen inequality

In this section, we use $I \subset \mathbb{R}$ as an interval, since we use J for an alternative symbol. Our purpose of this section is to give a better bounds on operator Jensen inequality. It is known that [3], if $f : I \to \mathbb{R}$ is a convex function, $A_1, \ldots, A_n$ self-adjoint operators with spectra contained in I, and $w_1, \ldots, w_n$ positive numbers such that $\sum_{i=1}^{n} w_i = 1$, then

$$f\left(\sum_{i=1}^{n} w_i \langle A_i x, x\rangle\right) \leq \sum_{i=1}^{n} w_i \langle f(A_i)x, x\rangle, \tag{5.5.1}$$

where $x \in \mathcal{H}$ with $\|x\| = 1$. In the following theorem, we make a refinement of the inequality (5.5.1).

Theorem 5.5.1 ([179]). *Let $f : I \to \mathbb{R}$ be a convex function, $A_1, \dots, A_n$ self-adjoint operators with spectra contained in I, and $w_1, \dots, w_n$ positive numbers such that $\sum_{i=1}^{n} w_i = 1$. Assume $J \subsetneq \{1, 2, \dots, n\}$ and $J^c = \{1, 2, \dots, n\} \backslash J$, $\omega_J = \sum_{i \in J} w_i$, $\omega_{J^c} = 1 - \sum_{i \in J} w_i$. Then for any unit vector $x \in \mathcal{H}$,*

$$f\left(\sum_{i=1}^{n} w_i \langle A_i x, x\rangle\right) \le \Psi(f, \mathbb{A}, J, J^c) \le \sum_{i=1}^{n} w_i \langle f(A_i)x, x\rangle, \tag{5.5.2}$$

where

$$\Psi(f, \mathbb{A}, J, J^c) = \omega_J f\left(\frac{1}{\omega_J} \sum_{i \in J} w_i \langle A_i x, x\rangle\right) + \omega_{J^c} f\left(\frac{1}{\omega_{J^c}} \sum_{i \in J^c} w_i \langle A_i x, x\rangle\right).$$

The inequality (5.5.2) *reverses if the function f is concave on I.*

Proof. We can replace x_i by $\langle A_i x, x\rangle$ where $x \in \mathcal{H}$ and $\|x\| = 1$, in (5.0.3). Hence, by using [97, Theorem 1.2], we can immediately infer that

$$f\left(\frac{1}{W_n} \sum_{i=1}^{n} w_i \langle A_i x, x\rangle\right) \le \frac{1}{W_n} \sum_{i=1}^{n} w_i f(\langle A_i x, x\rangle) \le \frac{1}{W_n} \sum_{i=1}^{n} w_i \langle f(A_i)x, x\rangle, \tag{5.5.3}$$

where $W_n = \sum_{i=1}^{n} w_i$. Now a simple calculation shows that

$$\begin{aligned} \sum_{i=1}^{n} w_i \langle f(A_i)x, x\rangle &= \sum_{i \in J} w_i \langle f(A_i)x, x\rangle + \sum_{i \in J^c} w_i \langle f(A_i)x, x\rangle \\ &= \omega_J\left(\frac{1}{\omega_J} \sum_{i \in J} w_i \langle f(A_i)x, x\rangle\right) + \omega_{J^c}\left(\frac{1}{\omega_{J^c}} \sum_{i \in J^c} w_i \langle f(A_i)x, x\rangle\right) \\ &\ge \omega_J f\left(\frac{1}{\omega_J} \sum_{i \in J} w_i \langle A_i x, x\rangle\right) + \omega_{J^c} f\left(\frac{1}{\omega_{J^c}} \sum_{i \in J^c} w_i \langle A_i x, x\rangle\right) = \Psi(f, \mathbb{A}, J, J^c), \end{aligned} \tag{5.5.4}$$

where we used the inequality (5.5.3). On the other hand,

$$\begin{aligned} \Psi(f, \mathbb{A}, J, J^c) &= \omega_J f\left(\frac{1}{\omega_J} \sum_{i \in J} w_i \langle A_i x, x\rangle\right) + \omega_{J^c} f\left(\frac{1}{\omega_{J^c}} \sum_{i \in J^c} w_i \langle A_i x, x\rangle\right) \\ &\ge f\left(\omega_J\left(\frac{1}{\omega_J} \sum_{i \in J} w_i \langle A_i x, x\rangle\right) + \omega_{J^c}\left(\frac{1}{\omega_{J^c}} \sum_{i \in J^c} w_i \langle A_i x, x\rangle\right)\right) = f\left(\sum_{i=1}^{n} w_i \langle A_i x, x\rangle\right). \end{aligned} \tag{5.5.5}$$

In the above computation, we have used the assumption that f is a convex function. Thus, relation (5.5.4), together with inequality (5.5.5), yields the inequality (5.5.2). □

The following refinements of the arithmetic–geometric-harmonic mean inequality are of interest.

Corollary 5.5.1. *Let $a_1, \ldots, a_n$ be positive numbers and let $\{w_i\}, J, J^c$ be as in Theorem 5.5.1. Then we have*

$$\begin{aligned}\left(\sum_{i=1}^n w_i a_i^{-1}\right)^{-1} &\le \left(\frac{1}{\omega_J}\sum_{i\in J} w_i a_i^{-1}\right)^{-\omega_J}\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i a_i^{-1}\right)^{-\omega_{J^c}}\\ &\le \prod_{i=1}^n a_i^{w_i} \le \left(\frac{1}{\omega_J}\sum_{i\in J} w_i a_i\right)^{\omega_J}\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i a_i\right)^{\omega_{J^c}} \le \sum_{i=1}^n w_i a_i\end{aligned}$$

and

$$\begin{aligned}\left(\sum_{i=1}^n w_i a_i^{-1}\right)^{-1} &\le \left(\omega_J \prod_{i\in J} a_i^{-\frac{w_i}{\omega_J}} + \omega_{J^c}\prod_{i\in J^c} a_i^{-\frac{w_i}{\omega_{J^c}}}\right)^{-1}\\ &\le \prod_{i=1}^n a_i^{w_i} \le \omega_J \prod_{i\in J} a_i^{\frac{w_i}{\omega_J}} + \omega_{J^c}\prod_{i\in J^c} a_i^{\frac{w_i}{\omega_{J^c}}} \le \sum_{i=1}^n w_i a_i.\end{aligned}$$

By virtue of Theorem 5.5.1, we have the following result.

Corollary 5.5.2. *Let $f : I \to \mathbb{R}$ be a nonnegative increasing convex function, $A_1, \ldots, A_n$ positive operators with spectra contained in I, and let $\{w_i\}, J, J^c$ be as in Theorem 5.5.1. Then*

$$f\left(\left\|\sum_{i=1}^n w_i A_i\right\|\right) \le \omega_J f\left(\frac{1}{\omega_J}\left\|\sum_{i=1}^n w_i A_i\right\|\right) + \omega_{J^c} f\left(\frac{1}{\omega_{J^c}}\left\|\sum_{i=1}^n w_i A_i\right\|\right) \le \left\|\sum_{i=1}^n w_i f(A_i)\right\|. \tag{5.5.6}$$

The inequality (5.5.6) *reverses if the function f is nonnegative increasing concave on I.*

Proof. On account of assumptions, we can write

$$\begin{aligned}\sup_{\|x\|=1} f\left(\sum_{i=1}^n w_i \langle A_i x, x\rangle\right) &= f\left(\sup_{\|x\|=1}\left\langle \sum_{i=1}^n w_i A_i x, x\right\rangle\right) = f\left(\left\|\sum_{i=1}^n w_i A_i\right\|\right)\\ &\le \omega_J f\left(\frac{1}{\omega_J}\left\|\sum_{i\in J} w_i A_i\right\|\right) + \omega_{J^c} f\left(\frac{1}{\omega_{J^c}}\left\|\sum_{i\in J^c} w_i A_i\right\|\right)\\ &\le \sup_{\|x\|=1}\left\langle \sum_{i=1}^n w_i f(A_i) x, x\right\rangle = \left\|\sum_{i=1}^n w_i f(A_i)\right\|.\end{aligned}$$

This completes the proof. □

We give a remark on Corollary 5.5.2.

Remark 5.5.1. Let $A_1,\dots,A_n$ be positive operators and let $\{w_i\},J,J^c$ be as in Theorem 5.5.1. Then for any $r \ge 1$,

$$\left\|\sum_{i=1}^n w_iA_i\right\|^r \le \omega_J\left(\frac{1}{\omega_J}\left\|\sum_{i=1}^n w_iA_i\right\|\right)^r + \omega_{J^c}\left(\frac{1}{\omega_{J^c}}\left\|\sum_{i=1}^n w_iA_i\right\|\right)^r \le \left\|\sum_{i=1}^n w_iA_i^r\right\|. \tag{5.5.7}$$

For $0 < r \le 1$, the reverse inequalities hold. If every operator A_i is strictly positive, (5.5.7) is also true for $r < 0$.

The multiple version of the inequality (5.0.5) is proved in [171, Theorem 1] as follows: Let $f : I \to \mathbb{R}$ be an operator convex, $\Phi_1,\dots,\Phi_n$ normalized positive linear maps from $\mathbb{B}(\mathcal{H})$ to $\mathbb{B}(\mathcal{K})$, $A_1,\dots,A_n$ self-adjoint operators with spectra contained in I, and $w_1,\dots,w_n$ positive numbers such that $\sum_{i=1}^n w_i = 1$, then we have

$$f\left(\sum_{i=1}^n w_i\Phi_i(A_i)\right) \le \sum_{i=1}^n w_i f(\Phi_i(A_i)). \tag{5.5.8}$$

The following is a refinement of (5.5.8).

Theorem 5.5.2 ([179]). *Let $f : I \to \mathbb{R}$ be an operator convex, $\Phi_1,\dots,\Phi_n$ normalized positive linear maps from $\mathbb{B}(\mathcal{H})$ to $\mathbb{B}(\mathcal{K})$, $A_1,\dots,A_n$ self-adjoint operators with spectra contained in I, and let $\{w_i\},J,J^c$ be as in Theorem* 5.5.1. *Then we have*

$$f\left(\sum_{i=1}^n w_i\Phi_i(A_i)\right) \le \Delta(f,\mathbb{A},J,J^c) \le \sum_{i=1}^n w_i\Phi_i(f(A_i)), \tag{5.5.9}$$

where

$$\Delta(f,\mathbb{A},J,J^c) = \omega_J f\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i)\right) + \omega_{J^c} f\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i)\right).$$

The inequality (5.5.9) *reverses if the function f is operator concave on I.*

Proof. Note that

$$f\left(\frac{1}{W_n}\sum_{i=1}^n w_i\Phi_i(A_i)\right) \le \frac{1}{W_n}\sum_{i=1}^n w_i\Phi_i(f(A_i)), \tag{5.5.10}$$

where $W_n = \sum_{i=1}^n w_i$. By employing the inequality (5.5.10), we have

$$\begin{aligned}\sum_{i=1}^n w_i\Phi_i(f(A_i)) &= \sum_{i\in J} w_i\Phi_i(f(A_i)) + \sum_{i\in J^c} w_i\Phi_i(f(A_i))\\ &= \omega_J\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(f(A_i))\right) + \omega_{J^c}\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(f(A_i))\right)\end{aligned}$$

$$\geq \omega_J f\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i)\right) + \omega_{J^c} f\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i)\right) = \Delta(f, \mathbb{A}, J, J^c). \tag{5.5.11}$$

On the other hand, since f is an operator convex we get

$$\begin{aligned}\Delta(f, \mathbb{A}, J, J^c) &= \omega_J f\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i)\right) + \omega_{J^c} f\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i)\right)\\ &\geq f\left(\omega_J\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i)\right) + \omega_{J^c}\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i)\right)\right) = f\left(\sum_{i=1}^{n} w_i\Phi_i(A_i)\right).\end{aligned} \tag{5.5.12}$$

Combining the two inequalities (5.5.11) and (5.5.12), we have the desired inequality. □

A special case of (5.5.9) is the following statement.

Remark 5.5.2. Let $\Phi_1, \ldots, \Phi_n$ be normalized positive linear maps from $\mathbb{B}(\mathcal{H})$ to $\mathbb{B}(\mathcal{K})$, $A_1, \ldots, A_n$ self-adjoint operators with spectra contained in I, and let $\{w_i\}, J, J^c$ be as in Theorem 5.5.1. Then for any $r \in [-1, 0] \cup [1, 2]$,

$$\left(\sum_{i=1}^{n} w_i\Phi_i(A_i)\right)^r \leq \omega_J\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i)\right)^r + \omega_{J^c}\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i)\right)^r \leq \sum_{i=1}^{n} w_i\Phi_i(A_i^r).$$

For $r \in [0, 1]$, the reverse inequalities hold.

Corollary 5.5.3. *Let $\Phi_1, \ldots, \Phi_n$ be normalized positive linear maps from $\mathbb{B}(\mathcal{H})$ to $\mathbb{B}(\mathcal{K})$, $A_1, \ldots, A_n$ self-adjoint operators with spectra contained in I, and let $\{w_i\}, J, J^c$ be as in Theorem* 5.5.1. *Then for any $r \geq 1$ and every unitarily invariant norm,*

$$\begin{aligned}\left\|\left(\sum_{i=1}^{n} w_i\Phi_i(A_i)\right)^r\right\|_u &\leq \left\|\left(\omega_J\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i^r)\right)^{\frac{1}{r}} + \omega_{J^c}\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i^r)\right)^{\frac{1}{r}}\right)^r\right\|_u\\ &\leq \left\|\sum_{i=1}^{n} w_i\Phi_i(A_i^r)\right\|_u.\end{aligned} \tag{5.5.13}$$

In particular,

$$\begin{aligned}\left\|\left(\sum_{i=1}^{n} w_i X_i^* A_i X_i\right)^r\right\|_u &\leq \left\|\left(\omega_J\left(\frac{1}{\omega_J}\sum_{i\in J} w_i X_i^* A_i^r X_i\right)^{\frac{1}{r}} + \omega_{J^c}\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i X_i^* A_i^r X_i\right)^{\frac{1}{r}}\right)^r\right\|_u\\ &\leq \left\|\sum_{i=1}^{n} w_i X_i^* A_i^r X_i\right\|_u,\end{aligned} \tag{5.5.14}$$

where $X_1, \ldots, X_n$ are contractions with $\sum_{i=1}^{n} X_i^* X_i = I_{\mathcal{H}}$.

Proof. Of course, the inequality (5.5.14) is a direct consequence of inequality (5.5.13), so we prove (5.5.13). It follows from Remark 5.5.2 that

$$\begin{aligned}\left\|\sum_{i=1}^{n} w_i\Phi_i(A_i^{\frac{1}{r}})\right\|_u &\le \left\|\omega_J\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i)\right)^{\frac{1}{r}} + \omega_{J^c}\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i)\right)^{\frac{1}{r}}\right\|_u \\ &\le \left\|\left(\sum_{i=1}^{n} w_i\Phi_i(A_i)\right)^{\frac{1}{r}}\right\|_u\end{aligned}$$

for any $r \ge 1$. Replacing A_i by A_i^r, we get

$$\begin{aligned}\left\|\sum_{i=1}^{n} w_i\Phi_i(A_i)\right\|_u &\le \left\|\omega_J\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i^r)\right)^{\frac{1}{r}} + \omega_{J^c}\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i^r)\right)^{\frac{1}{r}}\right\|_u \\ &\le \left\|\left(\sum_{i=1}^{n} w_i\Phi_i(A_i^r)\right)^{\frac{1}{r}}\right\|_u. \end{aligned} \tag{5.5.15}$$

It is well known that $\|X\|_u^{(r)} = \||X|^r\|_u^{\frac{1}{r}}$ defines a unitarily invariant norm. So (5.5.15) implies

$$\begin{aligned}\left\|\left(\sum_{i=1}^{n} w_i\Phi_i(A_i)\right)^r\right\|_u &\le \left\|\left(\omega_J\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i^r)\right)^{\frac{1}{r}} + \omega_{J^c}\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i^r)\right)^{\frac{1}{r}}\right)^r\right\|_u \\ &\le \left\|\sum_{i=1}^{n} w_i\Phi_i(A_i^r)\right\|_u.\end{aligned}$$

The proof is complete. □

We know that (see the estimate in [136, equation (16)]) if σ is an operator mean in the Kubo–Ando sense and $A_i, B_i > 0$, then we have

$$\sum_{i=1}^{n} w_i(A_i\sigma B_i) \le \left(\sum_{i=1}^{n} w_iA_i\right)\sigma\left(\sum_{i=1}^{n} w_iB_i\right). \tag{5.5.16}$$

The following corollary can be regarded as a refinement and generalization of the inequality (5.5.16).

Corollary 5.5.4. *Let σ be an operator mean, $\Phi_1,\dots,\Phi_n$ normalized positive linear maps from $\mathbb{B}(\mathcal{H})$ to $\mathbb{B}(\mathcal{K})$, $A_1,\dots,A_n$, $B_1,\dots,B_n$ strictly positive operators with spectra contained in I, and let $\{w_i\}, J, J^c$ be as in Theorem* 5.5.1. *Then we have*

$$\begin{aligned}\sum_{i=1}^{n} w_i\Phi_i(A_i\sigma B_i) &\le \left(\sum_{i\in J} w_i\Phi_i(A_i)\right)\sigma\left(\sum_{i\in J} w_i\Phi_i(B_i)\right) + \left(\sum_{i\in J^c} w_i\Phi_i(A_i)\right)\sigma\left(\sum_{i\in J^c} w_i\Phi_i(B_i)\right) \\ &\le \left(\sum_{i=1}^{n} w_i\Phi_i(A_i)\right)\sigma\left(\sum_{i=1}^{n} w_i\Phi_i(B_i)\right).\end{aligned}$$

Proof. If $F(\cdot,\cdot)$ is a jointly operator concave, then Theorem 5.5.2 implies

$$\begin{aligned}
&\sum_{i=1}^{n} w_i\Phi_i(F(A_i,B_i)) \\
&\quad\le \omega_J F\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i), \frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(B_i)\right) + \omega_{J^c} F\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i), \frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(B_i)\right) \\
&\quad\le F\left(\sum_{i=1}^{n} w_i\Phi_i(A_i), \sum_{i=1}^{n} w_i\Phi_i(B_i)\right).
\end{aligned} \tag{5.5.17}$$

It is well known that $F(A,B) = A\sigma B$ is jointly concave, so it follows from (5.5.17) that

$$\begin{aligned}
\sum_{i=1}^{n} w_i\Phi_i(A_i\sigma B_i) &\le \omega_J\left(\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i)\right)\sigma\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(B_i)\right)\right) \\
&\quad + \omega_{J^c}\left(\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i)\right)\sigma\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(B_i)\right)\right) \\
&= \left(\sum_{i\in J} w_i\Phi_i(A_i)\right)\sigma\left(\sum_{i\in J} w_i\Phi_i(B_i)\right) + \left(\sum_{i\in J^c} w_i\Phi_i(A_i)\right)\sigma\left(\sum_{i\in J^c} w_i\Phi_i(B_i)\right) \\
&\le \left(\sum_{i=1}^{n} w_i\Phi_i(A_i)\right)\sigma\left(\sum_{i=1}^{n} w_i\Phi_i(B_i)\right),
\end{aligned}$$

thanks to the homogeneity of operator means. Hence the proof is completed. □

By setting $\sigma = \sharp_v$, ($v \in [0,1]$) and $\Phi_i(X_i) = X_i$ ($i = 1,\dots,n$) in Corollary 5.5.4, we improve the weighted operator Hölder and Cauchy inequalities in the following way.

Corollary 5.5.5. *Let $\Phi_1,\dots,\Phi_n$ be normalized positive linear maps from $\mathbb{B}(\mathcal{H})$ to $\mathbb{B}(\mathcal{K})$, $A_1,\dots,A_n$, $B_1,\dots,B_n$ strictly positive operators with spectra contained in I, and let $\{w_i\}, J, J^c$ be as in Theorem* 5.5.1. *Then for any $v \in [0,1]$,*

$$\begin{aligned}
\sum_{i=1}^{n} w_i(A_i\sharp_v B_i) &\le \left(\sum_{i\in J} w_iA_i\right)\sharp_v\left(\sum_{i\in J} w_iB_i\right) + \left(\sum_{i\in J^c} w_iA_i\right)\sharp_v\left(\sum_{i\in J^c} w_iB_i\right) \\
&\le \left(\sum_{i=1}^{n} w_iA_i\right)\sharp_v\left(\sum_{i=1}^{n} w_iB_i\right).
\end{aligned}$$

In particular,

$$\begin{aligned}
\sum_{i=1}^{n} w_i(A_i\sharp B_i) &\le \left(\sum_{i\in J} w_iA_i\right)\sharp\left(\sum_{i\in J} w_iB_i\right) + \left(\sum_{i\in J^c} w_iA_i\right)\sharp\left(\sum_{i\in J^c} w_iB_i\right) \\
&\le \left(\sum_{i=1}^{n} w_iA_i\right)\sharp\left(\sum_{i=1}^{n} w_iB_i\right).
\end{aligned}$$

The *operator perspective* enjoys the following property:

$$\mathcal{P}_f\big(\Phi(A)|\Phi(B)\big) \le \Phi\big(\mathcal{P}_f(A|B)\big).$$

This nice property has been proved by F. Hansen [101, 105]. Let us note that the perspective of an operator convex function is operator convex as a function of two variables (see [97, Theorem 2.2]). Taking into account above and applying Theorem 5.5.2, we get the following result.

Corollary 5.5.6. *Let $f : I \to \mathbb{R}$ be an operator convex, $\Phi_1, \dots, \Phi_n$ normalized positive linear maps from $\mathbb{B}(\mathcal{H})$ to $\mathbb{B}(\mathcal{K})$, $A_1, \dots, A_n$ self-adjoint operators with spectra contained in I, and let $\{w_i\}, J, J^c$ be as in Theorem* 5.5.1. *Then we have*

$$\begin{aligned}
&\mathcal{P}_f\left(\sum_{i=1}^n w_i\Phi_i(A_i) \,\Big|\, \sum_{i=1}^n w_i\Phi_i(B_i)\right)\\
&\le \omega_J\mathcal{P}_f\left(\frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(A_i) \,\Big|\, \frac{1}{\omega_J}\sum_{i\in J} w_i\Phi_i(B_i)\right) + \omega_{J^c}\mathcal{P}_f\left(\frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i) \,\Big|\, \frac{1}{\omega_{J^c}}\sum_{i\in J^c} w_i\Phi_i(A_i)\right)\\
&\le \sum_{i=1}^n w_i\Phi_i\big(\mathcal{P}_f(A_i|B_i)\big).
\end{aligned}$$

5.6 Operator inequalities via geometric convexity

Given a unitarily invariant norm $\|\cdot\|_u$ on $\mathbb{B}(\mathcal{H})$, for a finite dimensional Hilber space $\mathcal{H}$, the following **Hölder inequality for unitarily invariant norm** holds [127]:

$$\left\|A^{1-v}XB^v\right\|_u \le \|AX\|_u^{1-v}\|XB\|_u^v, \quad 0 \le v \le 1, \tag{5.6.1}$$

for positive operators A, B and an arbitrary $X \in \mathbb{B}(\mathcal{H})$.

In the paper [208], it was shown that the function $f(v) = \|A^{1-v}XB^v\|_u$ is log-convex on $\mathbb{R}$. This entails (5.6.1) and its reverse when $v \notin [0,1]$.

In this section, we present some applications of geometrically convex functions to operator inequalities. Recall that if J is a subinterval of $(0,\infty)$ and $f : J \to (0,\infty)$, then f is called **geometrically convex** [192] if

$$f(a^{1-v}b^v) \le f^{1-v}(a)f^v(b), \quad v \in [0,1]. \tag{5.6.2}$$

The power function x^r, $(x > 0,\ r \in \mathbb{R})$ satisfies the condition of geometrically convexity, with equality. The following functions are known as geometrically convex [192], $\exp, \sinh, \cosh$ on $(0,\infty)$, $\tan, \sec, \csc$ on $(0,\pi/2)$, $\arcsin$ on $(0,1]$, $-\log(1-x)$, $(1+x)/(1-x)$ on $(0,1)$ and $\sum\limits_{n=0}^{\infty} c_nx^n$, $(c_n \ge 0)$ on $(0,R_c)$, where R_c is a radius of convergence. It is easy to find that if the geometrically convex function $f : J \to (0,\infty)$ is decreasing, then f is also (usual) convex. The function $\sqrt{x+1}$ is an example such that

it is a geometrically convex and a (usual) concave. The concavity of $\sqrt{x+1}$ is trivial. The proof of the geometrically convexity of the function $\sqrt{x+1}$ is done by the use of the standard **Hölder inequality**:

$$\sum_{i=1}^{n} a_i b_i \le \left(\sum_{i=1}^{n} a_i^{\frac{1}{1-v}}\right)^{1-v} \left(\sum_{i=1}^{n} b_i^{\frac{1}{v}}\right)^{v}, \quad (0 < v < 1, a_i, b_i > 0)$$

with $n = 2$, $a_1 = 1$, $a_2 = a^{1-v}$, $b_1 = 1$, $b_2 = b^v$.

In addition, it is known that the function $f(x)$ is a geometrically convex if and only if the function $\log f(e^x)$ is (usual) convex; see, for example, [14].

Recall that the **numerical radius** $\omega(A)$ and usual **operator norm** $\|A\|$ of an operator A are defined by

$$\|A\| = \sup_{\|x\|=1} \|Ax\| \quad \text{and} \quad \omega(A) = \sup_{\|x\|=1} |\langle Ax, x\rangle|,$$

where $\|x\| = \sqrt{\langle x, x\rangle}$. Of course, $\omega(A)$ defines a norm on $\mathbb{B}(\mathcal{H})$ and for every $A \in \mathbb{B}(\mathcal{H})$, we have

$$\frac{1}{2}\|A\| \le \omega(A) \le \|A\|. \tag{5.6.3}$$

The second inequality in (5.6.3) has been improved considerably by F. Kittaneh in [129] as follows:

$$\omega(A) \le \frac{1}{2}\left\|(A^*A)^{\frac{1}{2}} + (AA^*)^{\frac{1}{2}}\right\| \le \frac{1}{2}\left(\|A\| + \|A^2\|^{\frac{1}{2}}\right) \le \|A\|. \tag{5.6.4}$$

On the other hand, S. S. Dragomir extended (5.6.4) to the product of two operators to the following form [41]:

$$\omega(B^*A)^r \le \frac{1}{2}\left\|(A^*A)^r + (B^*B)^r\right\|, \quad \text{for all } r \ge 1. \tag{5.6.5}$$

5.6.1 Scalar inequalities

In this subsection, we first present some results on (5.6.2). We begin by the following reverse of (5.6.2).

Lemma 5.6.1. *Let f be a geometrically convex on the interval J and $a, b \in J$. Then for any $v > 0$ or $v < -1$,*

$$f^{1+v}(a) f^{-v}(b) \le f(a^{1+v} b^{-v}). \tag{5.6.6}$$

The proof is not so difficult so that we omit it. We also present the following theorem without proof. See [215] for their proofs.

Theorem 5.6.1 ([215]). *Let f be a geometrically convex on the interval J and $a, b \in J$. Then for any $v \in [0,1]$,*

$$\left(\frac{f(\sqrt{ab})}{\sqrt{f(a)f(b)}}\right)^{2R} f^{1-v}(a)f^{v}(b) \le f(a^{1-v}b^{v}) \le \left(\frac{f(\sqrt{ab})}{\sqrt{f(a)f(b)}}\right)^{2r} f^{1-v}(a)f^{v}(b), \tag{5.6.7}$$

where $r = \min\{v, 1-v\}$ and $R = \max\{v, 1-v\}$.

The first inequality in (5.6.7) can be regarded as a reverse of (5.6.2). On the other hand, the second inequality in (5.6.7) provides a refinement of (5.6.2), since $\frac{f(\sqrt{ab})}{\sqrt{f(a)f(b)}} \le 1$.

We can extend (5.6.2) to the following form [192]:

$$f\left(\prod_{i=1}^{n} x_i^{w_i}\right) \le \prod_{i=1}^{n} f^{w_i}(x_i), \quad \sum_{i=1}^{n} w_i = 1. \tag{5.6.8}$$

We can extend (5.6.6) to the following.

Corollary 5.6.1. *Let $a, b_i, v_i \ge 0$ and let $v = \sum_{i=1}^{n} v_i$. If $f : (0,\infty) \to (0,\infty)$ is a geometrically convex, then*

$$f\left(a^{1+v}\prod_{i=1}^{n} b_i^{-v_i}\right) \ge f^{1+v}(a)\prod_{i=1}^{n} f^{-v_i}(b).$$

In the following, we aim to improve (5.6.8). To this end, we need the following simple lemma which can be proven by using (5.6.8).

Lemma 5.6.2. *Let f be a geometrically convex on the interval J and $x_1, \ldots, x_n \in J$, and $w_1, \ldots, w_n$ positive numbers with $\omega_n = \sum_{i=1}^{n} w_i$, then we have*

$$f\left(\left(\prod_{i=1}^{n} x_i^{w_i}\right)^{\frac{1}{\omega_n}}\right) \le \left(\prod_{i=1}^{n} f^{w_i}(x_i)\right)^{\frac{1}{\omega_n}}.$$

We also present the following theorem without proof. See [215] for the proof.

Theorem 5.6.2 ([215]). *Let f be a geometrically convex function on the interval I, $x_1, \ldots, x_n \in I$, and $w_1, \ldots, w_n$ positive numbers such that $\sum_{i=1}^{n} w_i = 1$. Assume $J \subsetneq \{1, 2, \ldots, n\}$ and $J^c = \{1, 2, \ldots, n\} \backslash J$, $\omega_J \equiv \sum_{i \in J} w_i$, $\omega_{J^c} = 1 - \sum_{i \in J} w_i$. Then we have*

$$f\left(\prod_{i=1}^{n} x_i^{w_i}\right) \le f\left(\left(\prod_{i\in J} x_i^{w_i}\right)^{\frac{1}{\omega_J}}\right)^{\omega_J} f\left(\left(\prod_{i\in J^c} x_i^{w_i}\right)^{\frac{1}{\omega_{J^c}}}\right)^{\omega_{J^c}} \le \prod_{i=1}^{n} f^{w_i}(x_i).$$

It is quite natural to consider the n-tuple version of Theorem 5.6.1.

Theorem 5.6.3 ([215]). *Let f be a geometrically convex on the interval J and $x_1, \ldots, x_n \in J$, and let $p_1, \ldots, p_n$ be nonnegative numbers with $\sum_{i=1}^{n} p_i = 1$. Then we have*

$$\left(\frac{f((\prod_{i=1}^n x_i)^{\frac{1}{n}})}{(\prod_{i=1}^n f(x_i))^{\frac{1}{n}}}\right)^{nR_n} \prod_{i=1}^n f^{p_i}(x_i) \le f\left(\prod_{i=1}^n x_i^{p_i}\right) \le \left(\frac{f((\prod_{i=1}^n x_i)^{\frac{1}{n}})}{(\prod_{i=1}^n f(x_i))^{\frac{1}{n}}}\right)^{nr_n} \prod_{i=1}^n f^{p_i}(x_i), \tag{5.6.9}$$

where $r_n = \min\{p_1, \ldots, p_n\}$ and $R_n = \max\{p_1, \ldots, p_n\}$.

Proof. We first prove the second inequality of (5.6.9). We may assume $r_n = p_k$ without loss of generality. For any $k = 1, \ldots, n$, we have

$$\begin{aligned}
\left(\frac{f((\prod_{i=1}^n x_i)^{\frac{1}{n}})}{(\prod_{i=1}^n f(x_i))^{\frac{1}{n}}}\right)^{nr_n} \prod_{i=1}^n f^{p_i}(x_i) &= f\left(\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}\right)^{np_k} \left(\prod_{i=1}^n f(x_i)^{\frac{p_i-p_k}{1-np_k}}\right)^{1-np_k} \\
&\ge f\left(\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}\right)^{np_k} f\left(\prod_{i=1}^n x_i^{\frac{p_i-p_k}{1-np_k}}\right)^{1-np_k} \\
&\ge f\left(\left(\prod_{i=1}^n x_i\right)^{\frac{np_k}{n}} \prod_{i=1}^n x_i^{(\frac{p_i-p_k}{1-np_k})(1-np_k)}\right) = f\left(\prod_{i=1}^n x_i^{p_i}\right).
\end{aligned}$$

In the above, the first inequality follows by (5.6.8) with $1 - np_k \ge 0$ and the second inequality follows by (5.6.2) with $a = \prod_{i=1}^n x_i^{\frac{1}{n}}$, $b = \prod_{i=1}^n x_i^{\frac{p_i-p_k}{1-np_k}}$, $1 - v = np_k$.

We also assume $R_n = p_l$ and for amy $l = 1, \ldots, n$, we have

$$\begin{aligned}
\left(\frac{f(\prod_{i=1}^n x_i^{p_i})}{\prod_{i=1}^n f(x_i)^{p_i}}\right)^{\frac{1}{np_l}} \left(\prod_{i=1}^n f(x_i)\right)^{\frac{1}{n}} &= f\left(\prod_{i=1}^n x_i^{p_i}\right)^{\frac{1}{np_l}} \left(\prod_{i=1}^n f(x_i)^{\frac{p_l-p_i}{np_l-1}}\right)^{\frac{np_l-1}{np_l}} \\
&\ge f\left(\prod_{i=1}^n x_i^{p_i}\right)^{\frac{1}{np_l}} f\left(\prod_{i=1}^n x_i^{\frac{p_l-p_i}{np_l-1}}\right)^{\frac{np_l-1}{np_l}} \\
&\ge f\left(\left(\prod_{i=1}^n x_i^{p_i}\right)^{\frac{1}{np_l}} \prod_{i=1}^n x_i^{(\frac{p_l-p_i}{np_l-1})(\frac{np_l-1}{np_l})}\right) = f\left(\prod_{i=1}^n x_i^{\frac{1}{n}}\right).
\end{aligned}$$

In the above, the first inequality follows by (5.6.8) with $\frac{np_l-1}{np_l} \ge 0$ and the first inequality follows by (5.6.2) with $a = \prod_{i=1}^n x_i^{p_i}$, $b = \prod_{i=1}^n x_i^{\frac{p_l-p_i}{np_l-1}}$, $1 - v = \frac{1}{np_l}$. Thus the first inequality of (5.6.9) was proven. □

We easily find that Theorem 5.6.3 recovers Theorem 5.6.1 when $n = 2$.

5.6.2 Numerical radius inequalities

In this subsection, we present an application of geometrically convex functions to *numerical radius inequalities*. Our first result in this direction is the general form of (5.6.5). For the rest of the subsection, geometrically convex functions are continuous and implicitly understood to be of the form $f : J \to (0,\infty)$, where J is a subinterval of $(0,\infty)$.

Theorem 5.6.4 ([215]). *Let $A, B \in \mathbb{B}(\mathcal{H})$ and f be an increasing convex function. Then*

$$f(\omega(B^*A)) \le \frac{1}{2}\|f(A^*A)+f(B^*B)\|. \tag{5.6.10}$$

Proof. We recall the following Jensen's-type inequality [97, Theorem 1.2]:

$$f(\langle Ax,x\rangle) \le \langle f(A)x,x\rangle, \tag{5.6.11}$$

for any unit vector $x \in \mathcal{H}$, where f is a convex function on J and A is a self-adjoint operator with spectrum contained in J. Now, let $x \in \mathcal{H}$ be a unit vector. We have

$$\begin{aligned} f(|\langle B^*Ax,x\rangle|) &= f(|\langle Ax,Bx\rangle|)\\ &\le f(\|Ax\|\|Bx\|) \quad \text{(by the Cauchy–Schwarz inequality)}\\ &= f\big(\sqrt{\langle Ax,Ax\rangle\langle Bx,Bx\rangle}\big)\\ &= f\big(\sqrt{\langle A^*Ax,x\rangle\langle B^*Bx,x\rangle}\big)\\ &\le f\Big(\frac{\langle A^*Ax,x\rangle+\langle B^*Bx,x\rangle}{2}\Big) \quad (f:\text{increasing and AM–GM inequality})\\ &\le \frac{1}{2}f(\langle A^*Ax,x\rangle)+\frac{1}{2}f(\langle B^*Bx,x\rangle) \quad (f:\text{convex})\\ &\le \frac{1}{2}(\langle f(A^*A)x,x\rangle+\langle f(B^*B)x,x\rangle) \quad \text{(by (5.6.11))}, \end{aligned}$$

where the last inequality follows form the arithmetic–geometric mean inequality. Thus, we have shown

$$f(|\langle B^*Ax,x\rangle|) \le \frac{1}{2}\langle f(A^*A)+f(B^*B)x,x\rangle.$$

By taking supremum over $x \in \mathcal{H}$ with $\|x\| = 1$ and the assumption that f is continuous increasing function, we get

$$\begin{aligned} f(\omega(B^*A)) &= f\Big(\sup_{\|x\|=1}|\langle B^*Ax,x\rangle|\Big) = \sup_{\|x\|=1} f(|\langle B^*Ax,x\rangle|)\\ &\le \frac{1}{2}\sup_{\|x\|=1}\langle f(A^*A)+f(B^*B)x,x\rangle = \frac{1}{2}\|f(A^*A)+f(B^*B)\|. \end{aligned}$$

Therefore, (5.6.10) holds. □

One can easily check that the function $f(t) = t^r$, $(r \geq 1)$ satisfies the assumptions in Theorem 5.6.4 on f. Thus the inequality (5.6.10) implies (5.6.5).

Corollary 5.6.2. *Let f as in Theorem 5.6.4 and let $A, B, X \in \mathbb{B}(\mathcal{H})$. Then*

$$f(\omega(A^* XB)) \leq \frac{1}{2} \| f(A^* |X^*|^{2\nu} A) + f(B^* |X|^{2(1-\nu)} B) \|, \quad (\nu \in [0,1]).$$

Proof. Let $X = U|X|$ be the polar decomposition of X. Then we have

$$f(\omega(A^* XB)) = f(\omega(A^* U|X|B)) = f(\omega((|X|^\nu U^* A)^* (|X|^{1-\nu} B)).$$

By substituting $B = |X|^\nu U^* A$ and $A = |X|^{1-\nu} B$ in Theorem 5.6.4, we get the desired inequality, noting that when $X = U|X|$, we have $|X^*| = U|X|U^*$ which implies $|X^*|^{2\nu} = U|X|^{2\nu} U^*$. □

The function $f(t) = t^r$ $(t > 0, r \geq 1)$ also satisfies the assumptions in Corollary 5.6.2. Thus Corollary 5.6.2 recovers the inequality given in [221]:

$$\omega^r(A^* XB) \leq \frac{1}{2} \| (A^* |X^*|^{2\nu} A)^r + (B^* |X|^{2(1-\nu)} B)^r \|, \quad (r \geq 1, \nu \in [0,1]).$$

Another interesting inequality for $f(\omega(AXB))$ may be obtained as follows. First, notice that if f is a convex function and $\alpha \leq 1$, it follows that

$$f(\alpha t) \leq \alpha f(t) + (1 - \alpha) f(0). \tag{5.6.12}$$

This follows by direct computations for the function $g(t) = f(\alpha t) - \alpha f(t)$.

For the coming results, we use the term **norm-contractive** to mean an operator X whose operator norm satisfies $\|X\| \leq 1$. The **norm-expansive** means $\|X\| \geq 1$.

Proposition 5.6.1. *Under the same assumptions as in Theorem 5.6.4, the following inequality holds for the norm-contractive X:*

$$f(\omega(A^* XB)) \leq \frac{\|X\|}{2} \| f(A^* A) + f(B^* B) \| + (1 - \|X\|) f(0).$$

In particular, if $f(0) = 0$, then we have

$$f(\omega(A^* XB)) \leq \frac{\|X\|}{2} \| f(A^* A) + f(B^* B) \|.$$

Proof. Proceeding as in Theorem 5.6.4 and noting (5.6.12), we have

$$\begin{aligned} f(|\langle B^* XAx, x \rangle|) &= f(|\langle XAx, Bx \rangle|) \leq f(\|XAx\| \|Bx\|) \\ &\leq f(\|X\| \|Ax\| \|Bx\|) \leq \|X\| f(\|Ax\| \|Bx\|) + (1 - \|X\|) f(0). \end{aligned}$$

Then an argument similar to Theorem 5.6.4 implies the desired inequality. □

In particular, if $f(t) = t^r$, we obtain the following extension of (5.6.5).

Corollary 5.6.3. *Let $A, B, X \in \mathbb{B}(\mathcal{H})$. If X is norm-contractive and $r \geq 1$, then we have*

$$\omega^r(A^*XB) \leq \frac{\|X\|^r}{2}\|(A^*A)^r + (B^*B)^r\|. \tag{5.6.13}$$

Proof. Notice that a direct application of Proposition 5.6.1 implies the weaker inequality

$$\omega^r(A^*XB) \leq \frac{\|X\|}{2}\|(A^*A)^r + (B^*B)^r\|.$$

However, noting the proof of Proposition 5.6.1 for the function $f(t) = t^r$, we have

$$\begin{aligned} f(|\langle B^*XAx, x\rangle|) &= f(|\langle XAx, Bx\rangle|) \leq f(\|XAx\|\|Bx\|) \\ &\leq f(\|X\|\|Ax\|\|Bx\|) = f(\|X\|)f(\|Ax\|\|Bx\|). \end{aligned}$$

Arguing as before implies the desired inequality. □

Corollary 5.6.3 recovers the inequality (5.6.5), when $X = I$. Our next target is to show similar inequalities for geometrically convex functions. For the purpose of our results, we remind the reader of the following inequality:

$$f(\langle Ax, x\rangle) \leq k(m, M, f)^{-1}\langle f(A)x, x\rangle, \tag{5.6.14}$$

valid for the concave function $f : [m, M] \to (0, \infty)$, the unit vector $x \in \mathcal{H}$ and the positive operator A satisfying $m \leq A \leq M$, for some positive scalars m, M. Here $k(m, M, f)$ is defined by

$$k(m, M, f) = \min\left\{\frac{1}{f(t)}\left(\frac{M-t}{M-m}f(m) + \frac{t-m}{M-m}f(M)\right) : t \in [m, M]\right\}. \tag{5.6.15}$$

Proposition 5.6.2. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be such that $0 < m \leq A, B \leq M$ and f be an increasing geometrically convex function. If f is also concave, then*

$$f(\omega(A^{\frac{1}{2}}XB^{\frac{1}{2}})) \leq \frac{\|X\|}{2K}\|f(A) + f(B)\|, \tag{5.6.16}$$

for the norm-expansive X, where $K := k(m, M, f)$.

Proof. Proceeding as in Proposition 5.6.1 and noting (5.6.14) and the inequality $f(\alpha t) \leq \alpha f(t)$ when f is concave and $\alpha \geq 1$, we obtain the desired inequality. □

In particular, the function $f(t) = t^r$, $(0 < r \leq 1)$ satisfies the conditions of Proposition 5.6.2. Further, noting that $f(\|X\|\|AX\|\|Bx\|) = f(\|X\|)f(\|Ax\|\|Bx\|)$, we obtain the inequality

$$\omega(A^{\frac{1}{2}}XB^{\frac{1}{2}}) \leq \left(\frac{\|X\|}{2K(h, r)}\right)^{\frac{1}{r}}\|A^r + B^r\|^{\frac{1}{r}},$$

for the positive operators A, B satisfying $0 < m \le A, B \le M$ and the norm-expansive X. The constant $K(h, r)$ is well known as a generalized Kantorovich constant given by the formula (2.0.8). See [97, Definition 2.2].

Theorem 5.6.5 ([215]). *Let $A, B \in \mathbb{B}(\mathcal{H})$ and f be an increasing convex function. Then for any $\alpha \in [0, 1]$,*

$$f\left(\left\|\frac{A+B}{2}\right\|\right) \le \frac{1}{4}\left(\left\|f(|A|^{2\alpha}) + f(|B|^{2\alpha})\right\| + \left\|f(|A^*|^{2(1-\alpha)}) + f(|B^*|^{2(1-\alpha)})\right\|\right) \tag{5.6.17}$$

and

$$f\left(w\left(\frac{A+B}{2}\right)\right) \le \frac{1}{4}\left\|f(|A|^{2\alpha}) + f(|A^*|^{2(1-\alpha)}) + f(|B|^{2\alpha}) + f(|B^*|^{2(1-\alpha)})\right\|. \tag{5.6.18}$$

Proof. Before proceeding, we recall the following useful inequality which is known in the literature as the **generalized mixed Schwarz inequality** (see, e. g., [128]):

$$|\langle Ax, y\rangle| \le \sqrt{\langle |A|^{2\alpha}x, x\rangle\langle |A^*|^{2(1-\alpha)}y, y\rangle}, \quad \alpha \in [0, 1], \tag{5.6.19}$$

where $A \in \mathbb{B}(\mathcal{H})$ and for any $x, y \in \mathcal{H}$. Let $x, y \in \mathcal{H}$ be unit vectors. We have

$$\begin{aligned}
&f\left(\frac{1}{2}|\langle (A+B)x, y\rangle|\right) \\
&\le f\left(\frac{1}{2}(|\langle Ax, y\rangle| + |\langle Bx, y\rangle|)\right) \le \frac{1}{2}(f(|\langle Ax, y\rangle|) + f(|\langle Bx, y\rangle|)) \\
&\le \frac{1}{2}\left(f\left(\sqrt{\langle |A|^{2\alpha}x, x\rangle\langle |A^*|^{2(1-\alpha)}y, y\rangle}\right) + f\left(\sqrt{\langle |B|^{2\alpha}x, x\rangle\langle |B^*|^{2(1-\alpha)}y, y\rangle}\right)\right) \\
&\le \frac{1}{2}\left(f\left(\frac{1}{2}(\langle |A|^{2\alpha}x, x\rangle + \langle |A^*|^{2(1-\alpha)}y, y\rangle)\right)\right) \\
&\quad + \frac{1}{2}\left(f\left(\frac{1}{2}(\langle |B|^{2\alpha}x, x\rangle + \langle |B^*|^{2(1-\alpha)}y, y\rangle)\right)\right) \\
&\le \frac{1}{4}(\langle f(|A|^{2\alpha}) + f(|B|^{2\alpha})x, x\rangle + \langle f(|A^*|^{2(1-\alpha)}) + f(|B^*|^{2(1-\alpha)})y, y\rangle),
\end{aligned}$$

where the first inequality follows from the triangle inequality and the fact that f is increasing, the second inequality follows from the convexity of f, the third inequality follows from (5.6.19), the fourth inequality follows that f is increasing and the arithmetic mean-geometric mean inequality, the last inequality follows from the convexity of f and (5.6.11).

Now, by taking supremum over $x, y \in \mathcal{H}$ with $\|x\| = \|y\| = 1$, we deduce the desired inequality (5.6.17). If we take $x = y$, and applying same procedure as above we get (5.6.18). □

The case $f(t) = t^r$, $(t > 0, r \ge 1)$ in Theorem 5.6.5 implies

$$\|A + B\|^r \le 2^{r-2}\left(\left\||A|^{2\alpha r} + |B|^{2\alpha r}\right\| + \left\||A^*|^{2(1-\alpha)r} + |B^*|^{2(1-\alpha)r}\right\|\right),$$

and

$$\omega^r(A+B) \le 2^{r-2}\left\| |A|^{2\alpha r} + |B|^{2\alpha r} + \left|A^*\right|^{2(1-\alpha)r} + \left|B^*\right|^{2(1-\alpha)r} \right\|.$$

Another observation led by Theorem 5.6.5 is the following extension; whose proof is identical to that of Theorem 5.6.5.

Corollary 5.6.4. *Let $A, B \in \mathbb{B}(\mathcal{H})$ and f be an increasing convex function. Then for any $\alpha, v \in [0,1]$,*

$$\begin{aligned} & f\left(\left\|(1-v)A + vB\right\|\right) \\ & \quad \le \frac{1}{2}\left(\left\|(1-v)f\left(|A|^{2\alpha}\right) + vf\left(|B|^{2\alpha}\right)\right\| + \left\|(1-v)f\left(\left|A^*\right|^{2(1-\alpha)}\right) + vf\left(\left|B^*\right|^{2(1-\alpha)}\right)\right\|\right). \end{aligned}$$

Next, we show the concave version of Theorem 5.6.5, which then entails new inequalities for $0 \le r \le 1$.

Theorem 5.6.6 ([215]). *Let $A, B \in \mathbb{B}(\mathcal{H}), \alpha \in [0,1]$ and f be an increasing geometrically convex function. Assume that, for positive scalars m, M,*

$$m \le |A|^{2\alpha}, \left|A^*\right|^{2(1-\alpha)}, |B|^{2\alpha}, \left|B^*\right|^{2(1-\alpha)} \le M.$$

If f is concave, then

$$f(\|A+B\|) \le \frac{1}{2K}\left(\left\|f\left(|A|^{2\alpha}\right) + f\left(|B|^{2\alpha}\right)\right\| + \left\|f\left(\left|A^*\right|^{2(1-\alpha)}\right) + f\left(\left|B^*\right|^{2(1-\alpha)}\right)\right\|\right) \tag{5.6.20}$$

and

$$f(w(A+B)) \le \frac{1}{2K}\left(\left\|f(|A|) + f\left(\left|A^*\right|\right) + f(|B|) + f\left(\left|B^*\right|\right)\right\|\right), \tag{5.6.21}$$

where $K := k(m, M, f)$.

Proof. The proof is similar to that of Theorem 5.6.5. However, we need to recall that a non-negative concave function f is subadditive, in the sense that $f(a+b) \le f(a) + f(b)$ and to recall (5.6.14). These will be needed to obtain the second and fifth inequalities below. All other inequalities follow as in Theorem 5.6.5. We have, for the unit vectors x, y,

$$\begin{aligned} & f\left(\left|\langle (A+B)x, y\rangle\right|\right) \\ & \quad \le f\left(\left(\left|\langle Ax, y\rangle\right| + \left|\langle Bx, y\rangle\right|\right)\right) \le f\left(\left|\langle Ax, y\rangle\right|\right) + f\left(\left|\langle Bx, y\rangle\right|\right) \\ & \quad \le f\left(\sqrt{\langle |A|^{2\alpha}x, x\rangle \langle \left|A^*\right|^{2(1-\alpha)}y, y\rangle}\right) + f\left(\sqrt{\langle |B|^{2\alpha}x, x\rangle \langle \left|B^*\right|^{2(1-\alpha)}y, y\rangle}\right) \\ & \quad \le \sqrt{f\left(\langle |A|^{2\alpha}x, x\rangle\right) f\left(\langle \left|A^*\right|^{2(1-\alpha)}y, y\rangle\right)} + \sqrt{f\left(\langle |B|^{2\alpha}x, x\rangle\right) f\left(\langle \left|B^*\right|^{2(1-\alpha)}y, y\rangle\right)} \\ & \quad \le K^{-1}\left(\sqrt{\langle f(|A|^{2\alpha})x, x\rangle \langle f\left(\left|A^*\right|^{2(1-\alpha)}\right)y, y\rangle} + \sqrt{\langle f(|B|^{2\alpha})x, x\rangle \langle f\left(\left|B^*\right|^{2(1-\alpha)}\right)y, y\rangle}\right) \\ & \quad \le \frac{1}{2K}\left(\langle f(|A|^{2\alpha}) + f(|B|^{2\alpha})x, x\rangle + \langle f\left(\left|A^*\right|^{2(1-\alpha)}\right) + f\left(\left|B^*\right|^{2(1-\alpha)}\right)y, y\rangle\right). \end{aligned}$$

Now, by taking supremum over unit vectors $x, y \in \mathcal{H}$, we deduce the desired inequalities. □

In particular, if $f(t) = t^r$, $0 \le r \le 1$, we obtain with $h = M/m$,

$$\|A+B\|^r \le \frac{1}{2k(h,r)}\left(\left\||A|^{2\alpha r} + |B|^{2\alpha r}\right\| + \left\||A^*|^{2(1-\alpha)r} + |B^*|^{2(1-\alpha)r}\right\|\right).$$

In the next result, we will use the concave-version of the inequality (5.6.11), where we have

$$f(\langle Ax, x\rangle) \ge \langle f(A)x, x\rangle, \tag{5.6.22}$$

when $x \in \mathcal{H}$ is a unit vector, $f : J \to \mathbb{R}$ is a concave function and A is self-adjoint with spectrum in J.

Theorem 5.6.7 ([215]). *Let $A \in \mathbb{B}(\mathcal{H})$, $0 \le \alpha \le 1$ and $f : J \to \mathbb{R}$ be an increasing geometrically convex function. If in addition f is convex, then we have*

$$f(\omega^2(A)) \le \left\|\alpha f(|A|^2) + (1-\alpha) f(|A^*|^2)\right\|.$$

Proof. Noting (5.6.19) and the monotonicity of f, we have for any unit vector $x \in \mathcal{H}$,

$$\begin{aligned}
f(|\langle Ax, x\rangle|^2) &\le f(\langle |A|^{2\alpha}x, x\rangle \langle |A^*|^{2(1-\alpha)}x, x\rangle) \\
&\le f(\langle |A|^2x, x\rangle^{\alpha} \langle |A^*|^2 x, x\rangle^{1-\alpha}) \quad (\text{by concavity of } t^{\alpha} \text{ and } t^{1-\alpha}) \\
&\le f^{\alpha}(\langle |A|^2x, x\rangle) f^{1-\alpha}(\langle |A^*|^2x, x\rangle) \quad (\text{by (5.6.2)}) \\
&\le \alpha f(\langle |A|^2x, x\rangle) + (1-\alpha) f(\langle |A^*|^2x, x\rangle) \quad (\text{by Young's inequality}) \\
&\le \alpha \langle f(|A|^2)x, x\rangle + (1-\alpha)\langle f(|A^*|^2)x, x\rangle \quad (\text{by (5.6.11)}) \\
&= \langle (\alpha f(|A|^2) + (1-\alpha) f(|A^*|^2))x, x\rangle \\
&\le \left\|f(|A|^2) + (1-\alpha) f(|A^*|^2)\right\|.
\end{aligned}$$

Therefore,

$$f(\omega^2(A)) = f\Big(\sup_{\|x\|=1} |\langle Ax, x\rangle|^2\Big) = \sup_{\|x\|=1} f(|\langle Ax, x\rangle|^2) \le \left\|f(|A|^2) + (1-\alpha) f(|A^*|^2)\right\|,$$

which completes the proof. □

Letting $f(t) = t^r$, $(t > 0, r \ge 1)$ in Theorem 5.6.7 implies

$$\omega^{2r}(A) \le \left\|\alpha |A|^{2r} + (1-\alpha)|A^*|^{2r}\right\|.$$

5.6.3 Matrix norms

In this subsection, we present some unitarily invariant norm inequalities on $M(n,\mathbb{C})$ via geometrically convexity. It is well known that the function $f(t) = \|A^tXB^t\|_u$ is log-convex on $\mathbb{R}$, for any unitarily invariant norm $\|\cdot\|_u$ and positive matrices A and B, [210]. In this subsection, we show that the function $f(t)$ defined above, is also geometrically convex, when A, B are expansive. In this context, we say that a matrix A is **expansive matrix** if $A \ge I$ and **contractive matrix** if $A \le I$.

Note that if $A, B \le I$, then for any $X \in M(n,\mathbb{C})$,

$$\|AXB\|_u \le \|A\|\|X\|_u\|B\| \le \|X\|_u,$$

by submultiplicativity of unitarily invariant norms. Therefore, if A, B are expansive and $\alpha \le \beta$, then for any X, $\|A^{\alpha-\beta}XB^{\alpha-\beta}\|_u \le \|X\|_u$, which gives, upon replacing X with $A^\beta XB^\beta$ $\|A^\alpha XB^\alpha\|_u \le \|A^\beta XB^\beta\|_u$. In particular, if $t_1, t_2 > 0$, then $\sqrt{t_1t_2} \le \frac{t_1+t_2}{2}$, and the above inequality implies

$$\|A^{\sqrt{t_1t_2}}XB^{\sqrt{t_1t_2}}\|_u \le \|A^{\frac{t_1+t_2}{2}}XB^{\frac{t_1+t_2}{2}}\|_u, \tag{5.6.23}$$

when A, B are expansive.

Theorem 5.6.8 ([215]). *If A, B are expansive and $X \in M(n,\mathbb{C})$ is arbitrary, then $f(t) = \|A^tXB^t\|_u$ is geometrically convex on $(0,\infty)$.*

Proof. Taking into account of (5.6.23), we obtain

$$f(\sqrt{t_1t_2}) = \|A^{\sqrt{t_1t_2}}XB^{\sqrt{t_1t_2}}\|_u \le \|A^{\frac{t_1+t_2}{2}}XB^{\frac{t_1+t_2}{2}}\|_u \le \|A^{t_1}XB^{t_1}\|_u^{\frac{1}{2}}\|A^{t_2}XB^{t_2}\|_u^{\frac{1}{2}},$$

where the last inequality follows from the well-known log-convexity of f. □

Corollary 5.6.5. *If A is expansive and B is contractive, then the function $f(t) = \|A^tXB^{1-t}\|_u$ is geometrically convex on $(0,\infty)$.*

Proof. Notice that if B is contractive, B^{-1} is expansive. Therefore, applying Theorem 5.6.8 with X replaced by XB and B replaced by B^{-1}, we obtain the result. □

Corollary 5.6.6. *Let A, B be expansive and let $X \in M(n,\mathbb{C})$ be arbitrary. Then we have $\|X\|_u \le \|AXB\|_u$ for any unitarily invariant norm $\|\cdot\|_u$.*

Proof. Let $f(t) = \|A^tXB^t\|_u$. By Theorem 5.6.8, f is geometrically convex on $[0,\infty)$. In particular, by letting $t_1 = 0, t_2 = 1$, in $f(\sqrt{t_1t_2}) \le \sqrt{f(t_1)f(t_2)}$, then we get the desired inequality. □

Corollary 5.6.7. *Let A, B be expansive and let $X \in M(n,\mathbb{C})$ be arbitrary. Then we have*

$$\|A^{t_1\sharp_v t_2}XB^{t_1\sharp_v t_2}\|_u \le \|A^{t_1}XB^{t_1}\|_u\sharp_v\|A^{t_2}XB^{t_2}\|_u$$

for $t_1, t_2 \ge 0$, $0 \le v \le 1$ and any unitarily invariant norm $\|\cdot\|_u$.

Proof. The proof follows the same guideline as in Theorem 5.6.8, where we have $t_1\sharp_v t_2 \le t_1\nabla_v t_2$. Then we have $f(t_1\sharp_v t_2) \le f(t_1\nabla_v t_2) \le f(t_1)\sharp_v f(t_2)$, where the last inequality follows from the log-convexity of f. □

Corollary 5.6.8. *Let $a_i, b_i \ge 1$, $t_1, t_2 \ge 0$ and $0 \le v \le 1$. Then we have*

$$\sum_{i=1}^{n}(a_ib_i)^{t_1\sharp_v t_2} \le \left(\sum_{i=1}^{n}(a_ib_i)^{t_1}\right)\sharp_v\left(\sum_{i=1}^{n}(a_ib_i)^{t_2}\right).$$

Proof. For the given parameters, define $A = \operatorname{diag}(a_i)$ and $B = \operatorname{diag}(b_i)$. Then clearly A and B are expansive. Applying Corollary 5.6.7 with X being the identity matrix implies the desired inequality. □

Notice that as a special case of Corollary 5.6.8, we obtain the **Cauchy–Schwarz-type inequality**

$$\sum_{i=1}^{n} a_ib_i \le \left(\sum_{i=1}^{n}(a_ib_i)^{t_1}\right)^{\frac{1}{2}}\left(\sum_{i=1}^{n}(a_ib_i)^{t_2}\right)^{\frac{1}{2}}$$

for the scalars $a_i, b_i \ge 1$ and $t_1, t_2 > 0$ satisfying $t_1t_2 = 1$.

5.7 Exponential inequalities for positive linear maps

In this section, we adopt the following notation. For a given function $f : [m, M] \to \mathbb{R}$, define

$$L[m, M, f](t) = a[m, M, f]t + b[m, M, f], \tag{5.7.1}$$

where

$$a[m, M, f] = \frac{f(M) - f(m)}{M - m} \quad \text{and} \quad b[m, M, f] = \frac{Mf(m) - mf(M)}{M - m}.$$

If no confusion arises, we will simply write $a[m, M, f] = a_f$ and $b[m, M, f] = b_f$. Also, for $t_0 \in (m, M)$, define

$$L'[t_0, f](t) = f(t_0) + f'(t_0)(t - t_0), \tag{5.7.2}$$

provided that $f'(t_0)$ exists. It is clear that for a convex function $f : [m, M] \to \mathbb{R}$, one has

$$L'[t_0, f](t) \le f(t) \le L[m, M, f](t), \tag{5.7.3}$$

while the inequalities are reversed for a concave function f.

Remark 5.7.1. Notice that if f is convex on an interval J containing $[m, M]$, then (5.7.3) is still valid for any $t_0 \in J$. That is, t_0 does not need to be in (m, M).

Now, if $f : [m, M] \to [0, \infty)$ is log-convex, we have the inequality $\log f(t) \leq L[m, M, \log f](t)$, which simply reads as follows:

$$f(t) \leq \left(f^{t-m}(M) f^{M-t}(m)\right)^{\frac{1}{M-m}} \leq L[m, M, f](t), \quad m \leq t \leq M, \tag{5.7.4}$$

where the second inequality is due to the arithmetic–geometric inequality. We refer the reader to [180] for some detailed discussion of (5.7.4). Another useful observations about log-convex functions is the following. If f is log-convex on $[m, M]$ and if $t_0 \in (m, M)$ is such that $f(t_0) \neq 0$, (5.7.3) implies

$$L'[t_0, g](t) \leq g, \quad \text{where } g = \log f.$$

Simplifying this inequality implies the following.

Lemma 5.7.1. *Let $f : [m, M] \to [0, \infty)$ be log-convex. If f is differentiable at $t_0 \in (m, M)$, then*

$$f(t) \geq f(t_0) \exp\left(\frac{f'(t_0)}{f(t_0)}(t - t_0)\right), \quad m \leq t \leq M.$$

In this section, we present several inequalities for log-convex functions based on the Mond–Pečarić method. In particular, we present inequalities that can be viewed as exponential inequalities for log-convex functions. More precisely, we present inequalities among the quantities

$$\Phi(f(A)), \quad f(\Phi(A)), \quad \Phi\left(\left(f^{A-m}(M) f^{M-A}(m)\right)^{\frac{1}{M-m}}\right)$$

and

$$\left(f^{\Phi(A)-m}(M) f^{M-\Phi(A)}(m)\right)^{\frac{1}{M-m}}.$$

An another interest in this section is to present inequalities for operator-like means when filtered through normalized positive linear maps. That is, it is known that for an operator mean σ, one has [172]

$$\Phi(A \sigma B) \leq \Phi(A) \sigma \Phi(B), \quad A, B \in \mathbb{B}^{+}(\mathcal{H}).$$

In particular, we show complementary inequalities for the geometric $\sharp_v$ and harmonic $!_v$ operator-like means for $v < 0$. When $v \notin [0, 1]$, these are not operator means. We adopt the notation $\natural_v$ and $\dagger_v$ for $v \notin [0, 1]$.

Proposition 5.7.1. *Let $f : [m, M] \to [0, \infty)$ be log-convex, $A \in \mathbb{B}(\mathcal{H})$ a self-adjoint operator with spectra contained in $[m, M]$ with $0 < m < M$ and Φ be a normalized positive*

linear map. Then we have

$$\frac{1}{K(m,M,f)}\Phi(f(A)) \le \frac{1}{K(m,M,f)}\Phi((f^{A-m}(M)f^{M-A}(m))^{\frac{1}{M-m}})$$
$$\le f(\Phi(A)) \le (f^{\Phi(A)-m}(M)f^{M-\Phi(A)}(m))^{\frac{1}{M-m}} \le K(m,M,f)\Phi(f(A)),$$

where $K(m,M,f)$ is defined in (5.2.11).

Proof. The first and the second inequalities follow from [180, Proposition 2.1] and the fact that $K(m,M,f) > 0$. So we have to prove the other inequalities. Applying a standard functional calculus argument for the operator $\Phi(A)$ in (5.7.4), we get

$$f(\Phi(A)) \le (f^{\Phi(A)-m}(M)f^{M-\Phi(A)}(m))^{\frac{1}{M-m}} \le L[m,M,f](\Phi(A)). \tag{5.7.5}$$

Following [164], we have for $\alpha > 0$,

$$L[m,M,f](\Phi(A)) - \alpha\Phi(f(A)) = a_f\Phi(A) + b_f - \alpha\Phi(f(A)) \le \beta,$$

where $\beta = \max\limits_{t\in[m,M]}\{a_f t + b_f - \alpha f(t)\}$. That is,

$$L[m,M,f](\Phi(A)) \le \alpha\Phi(f(A)) + \beta.$$

By setting $\beta = 0$, we obtain $\alpha = \max\limits_{t\in[m,M]}\{\frac{a_f t+b_f}{f(t)}\} =: K(m,M,f)$. With this choice of α, we have $L[m,M,f](\Phi(A)) \le K(m,M,f)\Phi(f(A))$, which, together with (5.7.5), complete the proof. □

Corollary 5.7.1. *Let $A \in \mathbb{B}(\mathcal{H})$ be a self-adjoint operator with spectra contained in $[m,M]$ with $0 < m < M$ and Φ be a normalized positive linear map. Then, for $t < 0$,*

$$\frac{1}{K(m,M,t)}\Phi(A^t) \le \frac{1}{K(m,M,t)}\Phi((M^{A-m}m^{M-A})^{\frac{t}{M-m}}) \le \Phi^t(A)$$
$$\le (M^{\Phi(A)-m}m^{M-\Phi(A)})^{\frac{t}{M-m}} \le K(m,M,t)\Phi(A^t),$$

where $K(m,M,t)$ is the generalized Kantorovich constant defined in (2.0.7).

Proof. The result follows immediately from Proposition 5.7.1, be letting $f(x) = x^t$, $(t < 0)$. □

Remark 5.7.2. Corollary 5.7.1 presents a refinement of the corresponding result in [161, Lemma 2].

As an another application of Proposition 5.7.1, we have the following bounds for operator means. To simplify our statement, we will adopt the following notation. For a given function $f : [m,M] \to [0,\infty)$ and two positive operators A and B satisfying $mA \le B \le MA$, we write

$$A\sigma_f B := A^{\frac{1}{2}}f(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{-\frac{1}{2}} \quad \text{and} \quad A\delta B = A^{-\frac{1}{2}}BA^{-\frac{1}{2}}.$$

Corollary 5.7.2. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two positive operators such that $mA \leq B \leq MA$ for some positive scalars m, M. Then, for any positive linear map Φ (not necessarily normalized) and any* log*-convex function $f : [m, M] \to [0, \infty)$,*

$$\frac{1}{K(m,M,f)}\Phi(A\sigma_f B) \leq \frac{1}{K(m,M,f)}\Phi(A^{\frac{1}{2}}(f^{A\delta B-m}(M)f^{M-A\delta B}(m))^{\frac{1}{M-m}}A^{\frac{1}{2}}) \leq \Phi(A)\sigma_f\Phi(B)$$
$$\leq \Phi^{\frac{1}{2}}(A)(f^{\Phi(A)\delta\Phi(B)-m}(M)f^{M-\Phi(A)\delta\Phi(B)}(m))^{\frac{1}{M-m}}\Phi^{\frac{1}{2}}(A)$$
$$\leq K(m,M,f)\Phi(A\sigma_f B).$$

Proof. From the assumption $mA \leq B \leq MA$, we have $m \leq A\delta B := A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \leq M$. Therefore, if f is log-convex on $[m, M]$, Proposition 5.7.1 implies

$$\frac{1}{K(m,M,f)}\psi(f(A\delta B)) \leq \frac{1}{K(m,M,f)}\psi((f^{A\delta B-m}(M)f^{M-A\delta B}(m))^{\frac{1}{M-m}}) \leq f(\psi(A\delta B))$$
$$\leq (f^{\psi(A\delta B)-m}(M)f^{M-\psi(A\delta B)}(m))^{\frac{1}{M-m}} \leq K(m,M,f)\psi(f(A\delta B)),$$

for any normalized positive linear map ψ. In particular, for the given Φ, define

$$\psi(X) = \Phi^{-\frac{1}{2}}(A)\Phi(A^{\frac{1}{2}}XA^{\frac{1}{2}})\Phi^{-\frac{1}{2}}(A).$$

Then ψ is a normalized positive linear map and the above inequalities imply, upon conjugating with $\Phi^{\frac{1}{2}}(A)$, the desired inequalities. □

In particular, Corollary 5.7.2 can be utilized to obtain the quasi-versions for the geometric and harmonic operator means, as follows.

Corollary 5.7.3. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two positive operators such that $mA \leq B \leq MA$ for some positive scalars m, M. Then, for any positive linear map Φ (not necessarily normalized) and for $v < 0$,*

$$\frac{1}{K(m,M,t)}\Phi(A\natural_t B) \leq \frac{1}{K(m,M,t)}\Phi(A\sigma_g B)$$
$$\leq \Phi(A)\natural_v\Phi(B) \leq \Phi(A)\sigma_g\Phi(B) \leq K(m,M,t)\Phi(A\natural_v B),$$

where $g(x) = (M^{x-m}m^{M-x})^{\frac{v}{M-m}}$ and $K(m, M, t)$ is as in Corollary 5.7.1.

Proof. Noting that the function $f(x) = x^v$ is log-convex on $[m, M]$ for $v < 0$, the result follows by the direct application of Corollary 5.7.2. □

Remark 5.7.3. If A, B are two positive operators, then we have

$$\Phi(A)\natural_v\Phi(B) \leq \Phi(A\natural_v B), \quad v \in [-1, 0).$$

Therefore, Corollary 5.7.3 can be regarded as an extension and a reverse for the above inequality, under the assumption $mA \leq B \leq MA$ with $M \geq m > 0$.

Corollary 5.7.4. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two positive operators such that $mA \le B \le MA$ for some positive scalars $1 \le m \le M$. Then, for any positive linear map Φ (not necessarily normalized) and for $v < 0$,*

$$\begin{aligned}\frac{1}{H(m,M,v)}\Phi(A\dagger_v B) &\le \frac{1}{H(m,M,t)}\Phi(A\sigma_g B) \le \Phi(A)\dagger_v\Phi(B) \le \Phi(A)\sigma_g\Phi(B)\\ &\le H(m,M,v)\Phi(A\dagger_v B),\end{aligned}$$

where $g(x) = ((1\dagger_v M)^{x-m}(1\dagger_v m)^{M-x})^{\frac{1}{M-m}}$ and

$$H(m,M,v) = \left\{(1-v)^2 + \frac{v}{mM}(2(1-v)\sqrt{mM} + v)\right\}(1\dagger_v m)(1\dagger_v M).$$

Proof. Noting that the function $f(x) = (1 - v + vx^{-1})^{-1}$ is log-convex on $[m, M]$ for $v < 0$, provided that $m \ge 1$, the result follows by the direct application of Corollary 5.7.2. □

We should remark that the mapping $v \mapsto H(m,M,v)$ is a decreasing function for $v < 0$. In particular,

$$1 = H(m,M,0) \le H(m,M,v) \le \lim_{v\to-\infty} H(m,M,v) = \frac{(\sqrt{mM}-1)^2}{(m-1)(M-1)}, \quad \forall v < 0.$$

Further, utilizing (5.7.4), we obtain the following. In this result and later in this section, we adopt the notation:

$$\alpha(f,t_0) = \frac{a_f}{f'(t_0)} \quad \text{and} \quad \beta(f,t_0) = a_f t_0 + b_f - \frac{a_f f(t_0)}{f'(t_0)}.$$

The following proposition gives a simplified special case of [162, Theorem 2.1].

Proposition 5.7.2. *Let $f : [m, M] \to [0, \infty)$ be convex, $A \in \mathbb{B}(\mathcal{H})$ a self-adjoint operator with spectra contained in $[m, M]$ with $0 < m < M$ and Φ be a normalized positive linear map. If f is either increasing or decreasing on $[m, M]$, then for any $t_0 \in (m, M)$,*

$$\Phi\big(f(A)\big) \le \alpha(f,t_0) f\big(\Phi(A)\big) + \beta(f,t_0), \tag{5.7.6}$$

and

$$f\big(\Phi(A)\big) \le \alpha(f,t_0)\Phi\big(f(A)\big) + \beta(f,t_0), \tag{5.7.7}$$

provided that $f'(t_0)$ exists and $f'(t_0) \ne 0$. Further, both inequalities are reversed if f is concave.

Proof. Notice first that f being either increasing or decreasing assures that $a_f f'(t_0) > 0$. Using a standard functional calculus in (5.7.3) with $t = A$ and applying Φ to both sides imply

$$f(t_0) + f'(t_0)\big(\Phi(A) - t_0\big) \le \Phi\big(f(A)\big) \le a_f\Phi(A) + b_f. \tag{5.7.8}$$

On the other hand, applying the functional calculus argument with $t = \Phi(A)$ implies

$$f(t_0) + f'(t_0)(\Phi(A) - t_0) \le f(\Phi(A)) \le a_f \Phi(A) + b_f. \tag{5.7.9}$$

Noting that a_f and $f'(t_0)$ have the same sign, both desired inequalities follow from (5.7.8) and (5.7.9). Now if f is concave, replacing f with $-f$ and noting linearity of Φ imply the desired inequalities for a concave function. □

As an application, we present the following result.

Corollary 5.7.5. *Let $A \in \mathbb{B}(\mathcal{H})$ be a self-adjoint operator with spectra contained in $[m, M]$ with $0 < m < M$. Then, for a normalized positive linear map Φ,*

$$\Phi(A^{-1}) \le \Phi^{-1}(A) + \left(\frac{1}{\sqrt{m}} - \frac{1}{\sqrt{M}}\right)^2 \tag{5.7.10}$$

and

$$\Phi(A^{-1}) \le \frac{(M+m)^2}{4mM}\Phi^{-1}(A). \tag{5.7.11}$$

Proof. Let $f(t) = t^{-1}$. Then f is convex and monotone on $[m, M]$. Letting $t_0 = \sqrt{mM} \in (m, M)$, direct calculations show that $\alpha(f, t_0) = 1, \beta(f, t_0) = (\frac{1}{\sqrt{m}} - \frac{1}{\sqrt{M}})^2$. Then inequality (5.7.6) implies the first inequality. The second inequality follows similarly by letting $t_0 = \frac{m+M}{2}$. □

Manipulating Proposition 5.7.2 implies several extensions for log-convex functions, as we shall see next. We will adopt the following constants in Theorem 5.7.1:

$a_h = a[m, M, h]$, $b_h = b[m, M, h]$, $\alpha = \alpha(h, t_0)$, $\beta = \beta(h, t_0)$ for $h(t) = (f^{t-m}(M) \times f^{M-t}(m))^{\frac{1}{M-m}}$ and $a_{h_1} = a[f^{M-m}(m), f^{M-m}(M), h_1]$, $b_{h_1} = b[f^{M-m}(m), f^{M-m}(M), h_1]$, $\alpha_1 = \alpha(h_1, t_1)$ and $\beta_1 = \beta(h_1, t_1)$ for $h_1(t) = t^{\frac{1}{M-m}}$. The first two inequalities of the next result should be compared with Proposition 5.7.1, where a reverse-type is presented now.

Theorem 5.7.1 ([214]). *Let $f : [m, M] \to [0, \infty)$ be* log-*convex, $A \in \mathbb{B}(\mathcal{H})$ a self-adjoint operator with spectra contained in $[m, M]$ with $0 < m < M$ and Φ be a normalized positive linear map. Then for any $t_0, t_1 \in (m, M)$,*

$$\begin{aligned} f(\Phi(A)) &\le (f^{\Phi(A)-m}(M) f^{M-\Phi(A)}(m))^{\frac{1}{M-m}} \le \alpha\Phi(f^{\frac{A-m}{M-m}}(M) f^{\frac{M-A}{M-m}}(m)) + \beta \\ &\le \begin{cases} \alpha(\Phi(f^{A-m}(M) f^{M-A}(m)))^{\frac{1}{M-m}} + \beta, & M - m \ge 1 \\ \alpha\alpha_1(\Phi(f^{A-m}(M) f^{M-A}(m)))^{\frac{1}{M-m}} + \alpha\beta_1 + \beta, & M - m < 1. \end{cases} \end{aligned}$$

Proof. For $h(t) = (f^{t-m}(M) f^{M-t}(m))^{\frac{1}{M-m}}$, we clearly see that h is convex and monotone on $[m, M]$. Notice that

$$
\begin{aligned}
f(\Phi(A)) &\le \left(f^{\Phi(A)-m}(M)f^{M-\Phi(A)}(m)\right)^{\frac{1}{M-m}} \quad \text{(by the third inequality of Proposition 5.7.1)}\\
&= h(\Phi(A)) \le \alpha(h,t_0)\Phi(h(A)) + \beta(h,t_0) \quad \text{(by (5.7.7))}\\
&= \alpha(h,t_0)\Phi\left(f^{\frac{A-m}{M-m}}(M)f^{\frac{M-A}{M-m}}(m)\right) + \beta(h,t_0).
\end{aligned}
$$

This proves the first two inequalities. Now, for the third inequality, assume that $M - m \ge 1$ and let $h_1(t) = t^{\frac{1}{M-m}}$. Then the second inequality can be viewed as

$$
f(\Phi(A)) \le \alpha\Phi\left(h_1(f^{A-m}(M)f^{M-A}(m))\right) + \beta. \tag{5.7.12}
$$

Since $M - m \ge 1$, it follows that h_1 is operator concave. Therefore, from (5.7.12) and (5.0.5) we can write

$$
f(\Phi(A)) \le \alpha\Phi\left(h_1(f^{A-m}(M)f^{M-A}(m))\right) + \beta \le \alpha h_1\left(\Phi(f^{A-m}(M)f^{M-A}(m))\right) + \beta,
$$

which is the desired inequality in the case $M - m \ge 1$.

Now, if $M - m < 1$, the function h_1 is convex and monotone. Therefore, taking into account of (5.7.12) and (5.7.6), we obtain

$$
\begin{aligned}
f(\Phi(A)) &\le \alpha\Phi\left(h_1(f^{A-m}(M)f^{M-A}(m))\right) + \beta\\
&\le \alpha\left(\alpha_1 h_1\left(\Phi(f^{A-m}(M)f^{M-A}(m))\right) + \beta_1\right) + \beta\\
&= \alpha\alpha_1\left(\Phi(f^{A-m}(M)f^{M-A}(m))\right)^{\frac{1}{M-m}} + \alpha\beta_1 + \beta,
\end{aligned}
$$

which completes the proof. □

For the same parameters as Theorem 5.7.1, we have the following comparison, too, in which the first two inequalities have been shown in Theorem 5.7.1.

Corollary 5.7.6. *Let $f : [m, M] \to [0, \infty)$ be* log-*convex, $A \in \mathbb{B}(\mathcal{H})$ a self-adjoint operator with spectra contained in $[m, M]$ with $0 < m < M$ and Φ be a normalized positive linear map. Then we have*

$$
\begin{aligned}
f(\Phi(A)) &\le \left(f^{\Phi(A)-m}(M)f^{M-\Phi(A)}(m)\right)^{\frac{1}{M-m}}\\
&\le \alpha\Phi\left(f^{\frac{A-m}{M-m}}(M)f^{\frac{M-A}{M-m}}(m)\right) + \beta \le \alpha K(m,M,f)\Phi(f(A)) + \beta.
\end{aligned}
$$

Proof. We prove the last inequality. Letting $\psi(X) = X$ be a normalized positive linear map and noting that Φ is order preserving, the fourth inequality of proposition 5.7.1 implies

$$
\begin{aligned}
\alpha\Phi\left(f^{\frac{A-m}{M-m}}(M)f^{\frac{M-A}{M-m}}(m)\right) + \beta &= \alpha\Phi\left(f^{\frac{\psi(A)-m}{M-m}}(M)f^{\frac{M-\psi(A)}{M-m}}(m)\right) + \beta\\
&\le \alpha K(m,M,f)\Phi(\psi(f(A))) + \beta = \alpha K(m,M,f)\Phi(f(A)) + \beta,
\end{aligned}
$$

which is the desired inequality. □

For the next result, the following constants will be used: $\hat{a}_h = a[m, M, \hat{h}]$, $\hat{b}_h = b[m, M, \hat{h}]$, $\hat{\alpha} = \hat{\alpha}(\hat{h}, t_0)$, $\hat{\beta} = \hat{\beta}(\hat{h}, t_0)$ for $\hat{h}(t) = f^{t-m}(M) f^{M-t}(m)$ and $a_{h_1} = a[f^{M-m}(m), f^{M-m}(M), h_1]$, $b_{h_1} = b[f^{M-m}(m), f^{M-m}(M), h_1]$, $\alpha_1 = \alpha(h_1, t_1)$ and $\beta_1 = \beta(h_1, t_1)$ for $h_1(t) = t^{\frac{1}{M-m}}$.

Theorem 5.7.2 ([214]). *Let $f : [m, M] \to [0, \infty)$ be* log*-convex, $t_0, t_1 \in (m, M)$, $A \in \mathbb{B}(\mathcal{H})$ a self-adjoint operator with spectra contained in $[m, M]$ with $0 < m < M$ and Φ be a normalized positive linear map. If $M - m \geq 1$, then we have*

$$\Phi(f(A)) \leq [\Phi(f^{A-m}(M) f^{M-A}(m))]^{\frac{1}{M-m}} \leq (\hat{\alpha} f^{\Phi(A)-m}(M) f^{M-\Phi(A)}(m) + \hat{\beta})^{\frac{1}{M-m}}.$$

On the other hand, if $M - m < 1$, then we have

$$\Phi(f(A)) \leq \alpha_1 [\Phi(f^{M-A}(m) f^{A-m}(M))]^{\frac{1}{M-m}} + \beta_1 \leq \alpha_1 [\hat{\alpha} f^{\Phi(A)-m}(M) f^{M-\Phi(A)}(m) + \hat{\beta}]^{\frac{1}{M-m}} + \beta_1.$$

Proof. Letting $h_1(t) = t^{\frac{1}{M-m}}$ and $\hat{h}(t) = f^{t-m}(M) f^{M-t}(m)$, we have

$$\begin{aligned}
\Phi(f(A)) &\leq \Phi(h_1(f^{A-m}(M) f^{M-A}(m))) \quad \text{(by the first inequality of Proposition 5.7.1)} \\
&\leq h_1(\Phi(f^{A-m}(M) f^{M-A}(m))) \quad \text{(since } h_1 \text{ is operator concave)} \\
&= [\Phi(f^{A-m}(M) f^{M-A}(m))]^{\frac{1}{M-m}} = [\Phi(\hat{h}(A))]^{\frac{1}{M-m}} \quad \text{(where } \hat{h}(t) = f^{t-m}(M) f^{M-t}(m)) \\
&\leq [\hat{\alpha} \hat{h}(\Phi(A)) + \beta]^{\frac{1}{M-m}} \quad \text{(by (5.7.6))} \\
&= (\hat{\alpha} f^{\Phi(A)-m}(M) f^{M-\Phi(A)}(m) + \hat{\beta})^{\frac{1}{M-m}},
\end{aligned}$$

which completes the proof for the case $M - m \geq 1$. Now if $M - m < 1$, we have

$$\begin{aligned}
\Phi(f(A)) &\leq \Phi(h_1(f^{A-m}(M) f^{M-A}(m))) \quad \text{(by the first inequality of Proposition 5.7.1)} \\
&\leq \alpha_1 h_1(\Phi(f^{A-m}(M) f^{M-A}(m))) + \beta_1 \quad \text{(by (5.7.6))} \\
&= \alpha_1 [\Phi(f^{A-m}(M) f^{M-A}(m))]^{\frac{1}{M-m}} + \beta_1 = \alpha_1 [\Phi(\hat{h}(A))]^{\frac{1}{M-m}} + \beta_1 \\
&\leq \alpha_1 [\hat{\alpha} \hat{h}(\Phi(A)) + \hat{\beta}]^{\frac{1}{M-m}} + \beta_1 \quad \text{(by (5.7.6))} \\
&= \alpha_1 [\hat{\alpha} f^{\Phi(A)-m}(M) f^{M-\Phi(A)}(m) + \hat{\beta}]^{\frac{1}{M-m}} + \beta_1,
\end{aligned}$$

which completes the proof. □

Remark 5.7.4. In both Theorems 5.7.1 and 5.7.2, the constants α and α_1 can be selected to be 1, as follows. Noting that the function h in both theorems is continuous on $[m, M]$ and differentiable on (m, M), the mean value theorem assures that $a_h = h'(t_0)$ for some $t_0 \in (m, M)$. This implies $\alpha = 1$, since we use the notation $\alpha = \alpha(h, t_0) = \frac{a_h}{h'(t_0)}$. A similar argument applies for h_1. These values of t_0 can be easily found. Moreover, one can find t_0 so that $\beta(h, t_0) = 0$, providing a multiplicative version. Since this is a direct application, we leave the tedious computations to the interested reader.

Utilizing Lemma 5.7.1, we obtain the following exponential inequality.

Proposition 5.7.3. *Let* $f : [m, M] \to [0, \infty)$ *be* log-*convex,* $t_0, t_1 \in (m, M)$, $A \in \mathbb{B}(\mathcal{H})$ *a self-adjoint operator with spectra contained in* $[m, M]$ *with* $0 < m < M$ *and* Φ *be a normalized positive linear map. Then we have*

$$\Phi(f(A)) \geq \frac{f(t_0)}{\alpha} \exp\left(\frac{f'(t_0)}{f(t_0)}(\Phi(A) - t_0)\right) - \frac{\beta}{\alpha} f(t_0),$$

and

$$f(\Phi(A)) \geq \frac{f(t_0)}{\alpha} \Phi\left(\exp\left(\frac{f'(t_0)}{f(t_0)}(A - t_0)\right)\right) - \frac{\beta}{\alpha} f(t_0),$$

where $\alpha = \alpha(k, t_1)$ *and* $\beta = \beta(k, t_1)$ *for* $k(t) = \exp(\frac{f'(t_0)}{f(t_0)}(t - t_0))$.

Proof. By Lemma 5.7.1, we have

$$f(t) \geq f(t_0) \exp\left(\frac{f'(t_0)}{f(t_0)}(t - t_0)\right), m \leq t \leq M.$$

A functional calculus argument applied to this inequality with $t = A$ implies

$$\begin{aligned}
\Phi(f(A)) &\geq f(t_0)\Phi\left(\exp\left(\frac{f'(t_0)}{f(t_0)}(A - t_0)\right)\right) \\
&= f(t_0)\Phi(k(A)) \geq f(t_0)\frac{k(\Phi(A)) - \beta}{\alpha} \quad \text{(by (5.7.7))} \\
&= \frac{f(t_0)}{\alpha} \exp\left(\frac{f'(t_0)}{f(t_0)}(\Phi(A) - t_0)\right) - \frac{\beta}{\alpha} f(t_0),
\end{aligned}$$

which completes the proof of the first inequality. The second inequality follows similarly using (5.7.6). □

The above results are all based on basic inequalities for convex functions. Therefore, refinements of convex functions' inequalities can be used to obtain sharper bounds. We give here some examples. In the paper [166], the following simple inequality was shown for the convex function $f : [m, M] \to \mathbb{R}$:

$$f(t) + \frac{2\min\{t - m, M - t\}}{M - m}\left(\frac{f(m) + f(M)}{2} - f\left(\frac{m + M}{2}\right)\right) \leq L[m, M, f](t). \tag{5.7.13}$$

This inequality can be used to obtain refinements of (5.7.6) and (5.7.7) as follows. First, we note that the function $t \mapsto t_{\min} := \frac{2\min\{t-m,M-t\}}{M-m}(\frac{f(m)+f(M)}{2} - f(\frac{m+M}{2}))$ is a continuous function. Further, noting that

$$\min\{t - m, M - t\} = \frac{M - m + |M + m - 2t|}{2},$$

one can apply a functional calculus argument on (5.7.13). With this convention, we will use the notation

$$A_{\min} := \frac{1}{M-m}\left(\frac{f(m)+f(M)}{2} - f\left(\frac{m+M}{2}\right)\right)(M - m + |M + m - A|).$$

The following is a refinement of Proposition 5.7.2. Since the proof of Proposition 5.7.4 is similar to that of Proposition 5.7.2 utilizing (5.7.13), we do not include it here.

Proposition 5.7.4. *Let $f : [m,M] \to [0,\infty)$ be convex, $A \in \mathbb{B}(\mathcal{H})$ a self-adjoint operator with spectra contained in $[m,M]$ with $0 < m < M$ and Φ be a normalized positive linear map. If f is either increasing or decreasing on $[m,M]$, then for any $t_0 \in (m,M)$,*

$$\Phi(f(A)) + \Phi(A_{\min}) \le \alpha(f,t_0)f(\Phi(A)) + \beta(f,t_0), \tag{5.7.14}$$

and

$$f(\Phi(A)) + (\Phi(A))_{\min} \le \alpha(f,t_0)\Phi(f(A)) + \beta(f,t_0), \tag{5.7.15}$$

provided that $f'(t_0)$ exists and $f'(t_0) \neq 0$.

Applying this refinement to the convex function $f(t) = t^{-1}$ implies refinements of both inequalities in Corollary 5.7.5 in the following.

Corollary 5.7.7. *Under the assumptions of Corollary 5.7.5, we have*

$$\Phi(A^{-1}) + \Phi(A_{\min}) \le \Phi^{-1}(A) + \left(\frac{1}{\sqrt{m}} - \frac{1}{\sqrt{M}}\right)^2 \tag{5.7.16}$$

and

$$\Phi(A^{-1}) + (\Phi(A))_{\min} \le \frac{(M+m)^2}{4mM}\Phi^{-1}(A). \tag{5.7.17}$$

Remark 5.7.5. The inequality (5.7.13) has been studied extensively in the literature, where numerous refining terms have been found. We refer the reader to [211, 212], where a comprehensive discussion has been made therein. These refinements can be used to obtain further refining terms for Proposition 5.7.2. Further, these refinements can be applied to log-convex functions, too. This refining approach leads to refinements of most inequalities presented in this book, where convexity is the key idea. We leave the detailed computations to the interested reader.

5.8 Jensen's inequality for strongly convex functions

Let $J \subset \mathbb{R}$ be an interval and c be a positive number. Following B. T. Polyak [203], a function $f : J \to \mathbb{R}$ is called **strongly convex with modulus** c if

$$f(vx + (1-v)y) \le vf(x) + (1-v)f(y) - cv(1-v)(x-y)^2, \tag{5.8.1}$$

for all $x, y \in J$ and $v \in [0,1]$. Obviously, every strongly convex function is convex. Observe also that, for instance, affine functions are not strongly convex. Since strong

convexity is strengthening the notion of convexity, some properties of strongly convex functions are just *stronger versions* of known properties of convex functions. For instance, a function $f : J \to \mathbb{R}$ is strongly convex with modulus c if and only if for every $x_0 \in J^o$ (the interior of J) there exists a number $l \in \mathbb{R}$ such that

$$c(x - x_0)^2 + l(x - x_0) + f(x_0) \le f(x), \quad x \in J. \tag{5.8.2}$$

In other words, f has a quadratic support at x_0. For differentiable functions f, f is strongly convex with modulus c if and only if

$$(f'(x) - f'(y))(x - y) \ge 2c(x - y)^2, \tag{5.8.3}$$

for each $x, y \in J$.

Consider a real valued function f defined on an interval J, $x_1, \dots, x_n \in J$ and $p_1, \dots, p_n \in [0, 1]$ with $\sum_{i=1}^{n} p_i = 1$. The **Jensen functional** is defined by

$$\mathcal{J}_n(f, \mathbf{x}, \mathbf{p}) = \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right).$$

According to [40, Theorem 1], if $\mathbf{x} = (x_1, \dots, x_n) \in I^n$, $\mathbf{p} = (p_1, \dots, p_n)$, $\mathbf{q} = (q_1, \dots, q_n)$ are nonnegative n-tuples satisfying $\sum_{i=1}^{n} p_i = 1$, $\sum_{i=1}^{n} q_i = 1$, $q_i > 0$, $i = 1, \dots, n$, then

$$\min_{1\le i\le n}\left\{\frac{p_i}{q_i}\right\}\mathcal{J}_n(f, \mathbf{x}, \mathbf{q}) \le \mathcal{J}_n(f, \mathbf{x}, \mathbf{p}) \le \max_{1\le i\le n}\left\{\frac{p_i}{q_i}\right\}\mathcal{J}_n(f, \mathbf{x}, \mathbf{q}).$$

A. McD. Mercer [157, Theorem 1.2] proved the following variant of Jensen inequality, to which we will refer as to the **Jensen–Mercer inequality**. If f is a convex function on $[m, M]$, then

$$f\left(M + m - \sum_{i=1}^{n} p_i x_i\right) \le f(M) + f(m) - \sum_{i=1}^{n} p_i f(x_i), \tag{5.8.4}$$

for all $x_i \in [m, M]$ and all $p_i \in [0, 1]$ ($i = 1, \dots, n$) with $\sum_{i=1}^{n} p_i = 1$. We refer the reader to [122, 154, 196] as a sample of the extensive use of this inequality in this field. The following lemma [158, Theorem 4] is the starting point for our discussion.

Lemma 5.8.1. *If $f : J \to \mathbb{R}$ is strongly convex with modulus c, then*

$$c\sum_{i=1}^{n} p_i(x_i - \overline{x})^2 \le \mathcal{J}_n(f, \mathbf{x}, \mathbf{p}), \tag{5.8.5}$$

for all $x_1, \dots, x_n \in J$, $p_1, \dots, p_n > 0$ with $\sum_{i=1}^{n} p_i = 1$ and $\overline{x} = \sum_{i=1}^{n} p_i x_i$.

In the paper [167], the following result has been given.

Theorem 5.8.1 ([167, Corollary 3]). *Let $f : J \to \mathbb{R}$ be a strongly convex function with modulus c, $\mathbf{x} = (x_1, \dots, x_n) \in J^n$, $\mathbf{p} = (p_1, \dots, p_n)$, $\mathbf{q} = (q_1, \dots, q_n)$ nonnegative n-tuples satisfying $\sum_{i=1}^n p_i = 1$, $\sum_{i=1}^n q_i = 1$, $q_i > 0$, $i = 1, \dots, n$. Then*

$$\begin{aligned}
& m\mathcal{J}_n(f, \mathbf{x}, \mathbf{q}) + c\left(\sum_{i=1}^n (p_i - mq_i)\left(x_i - \sum_{i=1}^n p_i x_i\right)^2 + m\left(\sum_{i=1}^n (p_i - q_i)x_i\right)^2\right) \\
& \le \mathcal{J}_n(f, \mathbf{x}, \mathbf{p}) \\
& \le M\mathcal{J}_n(f, \mathbf{x}, \mathbf{q}) - c\left(\sum_{i=1}^n (Mq_i - p_i)\left(x_i - \sum_{j=1}^n q_j x_j\right)^2 + \left(\sum_{i=1}^n (p_i - q_i)x_i\right)^2\right),
\end{aligned}$$

where $m := \min_{1\le i\le n}\{p_i/q_i\}$ and $M := \max_{1\le i\le n}\{p_i/q_i\}$.

Remark 5.8.1. Note that Lemma 5.8.1 also holds for functions defined on open convex subsets of an inner product space. See [197, Theorem 2]. Therefore, Theorem 5.8.1 can be also formulated and proved in such more general settings.

An interesting corollary can be deduced at this stage. We restrict ourselves to the case when $n = 2$.

Remark 5.8.2. According to J. B. Hiriart-Urruty and C. Lemaréchal [115, Proposition 1.1.2], the function $f : J \to \mathbb{R}$ is strongly convex with modulus c if and only if the function $g : J \to \mathbb{R}$ defined by $g(x) = f(x) - cx^2$ is convex. Hence the function $f : (0, 1] \to [0, \infty)$, $f(x) = -\log x$ is strongly convex with modulus $c = \frac{1}{2}$. Taking $p_1 = v$, $p_2 = 1 - v$, $q_1 = u$, $q_2 = 1 - u$ with $v, u \in [0, 1]$, $x_1 = a$, $x_2 = b$ and taking into account that $m = \min\{\frac{v}{u}, \frac{1-v}{1-u}\}$ and $M = \max\{\frac{v}{u}, \frac{1-v}{1-u}\}$, simple algebraic manipulations show that

$$\begin{aligned}
& \left(\frac{ua + (1-u)b}{a^u b^{1-u}}\right)^m \exp\left(\frac{(b-a)^2}{2}\left((v - mu)(v-1)^2 + v^2((1-v) - m(1-u)) + m(u-v)^2\right)\right) \\
& \quad \le \frac{va + (1-v)b}{a^v b^{1-v}} \\
& \quad \le \left(\frac{ua + (1-u)b}{a^u b^{1-u}}\right)^M \frac{1}{\exp\left(\frac{(b-a)^2}{2}\left((Mu - v)(u-1)^2 + u^2(M(1-u) - (1-v)) + (u-v)^2\right)\right)}.
\end{aligned}$$

Remark 5.8.2 admits the following important special case.

Corollary 5.8.1. *Let $a, b \in (0, 1]$, then*

$$\begin{aligned}
& K^r\left(\frac{a}{b}\right) \exp\left(\frac{(b-a)^2}{2}\left((v - r)(v-1)^2 + v^2((1-v) - r) + \frac{r}{2}(1-2v)^2\right)\right) \\
& \quad \le \frac{va + (1-v)b}{a^v b^{1-v}} \le \frac{K^R(\frac{a}{b})}{\exp\left(\frac{(b-a)^2}{8}\left((R - v) + (R - (1-v)) + (1-2v)^2\right)\right)},
\end{aligned}$$

where $r = \min\{v, 1 - v\}$, $R = \max\{v, 1 - v\}$ and $v \in [0, 1]$.

Proof. The result follows from Remark 5.8.2 by taking $u = \frac{1}{2}$. □

Remark 5.8.3. The following **Zuo–Liao inequality** holds:

$$K^r\left(\frac{a}{b}\right)a^v b^{1-v} \le va + (1-v)b \le K^R\left(\frac{a}{b}\right)a^v b^{1-v}, \tag{5.8.6}$$

where $a, b > 0$, $v \in [0,1]$, $r = \min\{v, 1-v\}$ and $R = \max\{v, 1-v\}$. Since $\exp(x) \ge 1$ for $x \ge 0$, Corollary 5.8.1 essentially gives a refinement of the inequalities in (5.8.6).

The following lemma plays an important role to prove Theorem 5.8.2.

Lemma 5.8.2. *If $f : J \to \mathbb{R}$ is strongly convex with modulus c, then*

$$f(x_1 + x_n - x_i) \le f(x_1) + f(x_n) - f(x_i) - 2cv_i(1-v_i)(x_1 - x_n)^2,$$

where $v_i \in [0,1]$, $x_1 = \min_{1\le i\le n} x_i$, $x_n = \max_{1\le i\le n} x_i$ and $x_i \in J$.

Proof. Let $y_i = x_1 + x_n - x_i$, $i = 1, 2, \ldots, n$. We may write $x_i = v_i x_1 + (1-v_i)x_n$ and $y_i = (1-v_i)x_1 + v_i x_n$ where $v_i \in [0,1]$.

Now, using simple calculations, we obtain

$$\begin{aligned}
f(y_i) &= f\big((1-v_i)x_1 + v_i x_n\big) \\
&\le (1-v_i)f(x_1) + v_i f(x_n) - cv_i(1-v_i)(x_1 - x_n)^2 \quad \text{(by (5.8.1))} \\
&= f(x_1) + f(x_n) - \big(v_i f(x_1) + (1-v_i)f(x_n)\big) - cv_i(1-v_i)(x_1 - x_n)^2 \\
&\le f(x_1) + f(x_n) - f\big(v_i x_1 + (1-v_i)x_n\big) - 2cv_i(1-v_i)(x_1 - x_n)^2 \quad \text{(by (5.8.1))} \\
&= f(x_1) + f(x_n) - f(x_i) - 2cv_i(1-v_i)(x_1 - x_n)^2,
\end{aligned}$$

which completes the proof. □

Lemma 5.8.2 follows also from the fact that strongly convex function is **strongly Wright-convex** (see, e. g., [190]). At this point, our aim is to present **Jensen–Mercer inequality** for strongly convex function.

Theorem 5.8.2 ([214]). *Let $f : J \to \mathbb{R}$ be a strongly convex with modulus c, then*

$$\begin{aligned}
&f\left(x_1 + x_n - \sum_{i=1}^n p_i x_i\right) \\
&\quad \le f(x_1) + f(x_n) - \sum_{i=1}^n p_i f(x_i) - c\left(2\sum_{i=1}^n p_i v_i(1-v_i)(x_1 - x_n)^2 + \sum_{i=1}^n p_i\left(x_i - \sum_{i=1}^n p_i x_i\right)^2\right),
\end{aligned}$$

where $\sum_{i=1}^n p_i = 1$, $v_i \in [0,1]$, $x_1 = \min_{1\le i\le n} x_i$, $x_n = \max_{1\le i\le n} x_i$ and $x_i \in J$.

Proof. A straightforward computation gives that

$$\begin{aligned}
f\left(x_1+x_n-\sum_{i=1}^n p_i x_i\right) &= f\left(\sum_{i=1}^n p_i(x_1+x_n-x_i)\right)\\
&\le \sum_{i=1}^n p_i f(x_1+x_n-x_i) - c\sum_{i=1}^n p_i\left(x_i-\sum_{i=1}^n p_i x_i\right)^2 \quad \text{(by Lemma 5.8.1)}\\
&\le f(x_1)+f(x_n)-\sum_{i=1}^n p_i f(x_i)\\
&\quad - c\left(2\sum_{i=1}^n p_i v_i(1-v_i)(x_1-x_n)^2+\sum_{i=1}^n p_i\left(x_i-\sum_{i=1}^n p_i x_i\right)^2\right) \quad \text{(by Lemma 5.8.2)},
\end{aligned}$$

as required. □

Remark 5.8.4. Based on Theorem 5.8.2, we obtain that

$$\begin{aligned}
&f\left(x_1+x_n-\sum_{i=1}^n p_i x_i\right)\\
&\quad\le f(x_1)+f(x_n)-\sum_{i=1}^n p_i f(x_i) - c\left(2\sum_{i=1}^n p_i v_i(1-v_i)(x_1-x_n)^2+\sum_{i=1}^n p_i\left(x_i-\sum_{i=1}^n p_i x_i\right)^2\right)\\
&\quad\le f(x_1)+f(x_n)-\sum_{i=1}^n p_i f(x_i).
\end{aligned}$$

Corollary 5.8.2. *Let us now define*

$$\widetilde{A} := x_1+x_n-\sum_{i=1}^n p_i x_i, \quad \widetilde{G} := \frac{x_1 x_n}{\prod_{i=1}^n x_i^{p_i}}.$$

As mentioned above, the function $f:(0,1]\to[0,\infty)$, $f(x)=-\log x$ *is strongly convex with modulus* $c=\frac{1}{2}$. *From Remark 5.8.4, we obtain* $-\ln\widetilde{A}\le -\ln\widetilde{G}-\frac{1}{2}\widetilde{M}\le -\ln\widetilde{G}$, *where* $\widetilde{M}=2\sum_{i=1}^n p_i v_i(1-v_i)(x_1-x_n)^2+\sum_{i=1}^n p_i\left(x_i-\sum_{i=1}^n p_i x_i\right)^2$. *We thus obtain*

$$\widetilde{G}\le e^{\frac{1}{2}\widetilde{M}}\widetilde{G}\le\widetilde{A}. \tag{5.8.7}$$

The celebrated **Hölder–McCarthy inequality** which is a special case of Theorem 5.1.1 asserts the following results.

Theorem 5.8.3 ([156]). *Let A be a positive operator on $\mathcal{H}$. If $x\in\mathcal{H}$ is a unit vector, then:*
(i) $\langle Ax,x\rangle^r \le \langle A^r x,x\rangle$ *for all* $r>1$.
(ii) $\langle Ax,x\rangle^r \ge \langle A^r x,x\rangle$ *for all* $0<r<1$.

Here, we give a more precise estimation than the inequality (5.1.1) for strongly convex functions with modulus c as follows:

Theorem 5.8.4 (Jensen operator inequality for strongly convex functions [214]). *Let $f : J \to \mathbb{R}$ be strongly convex with modulus c and differentiable on J^o. If A is a self-adjoint operator on the Hilbert space $\mathcal{H}$ with $Sp(A) \subset J^o$, then we have*

$$f(\langle Ax, x\rangle) \le \langle f(A)x, x\rangle - c(\langle A^2x, x\rangle - \langle Ax, x\rangle^2), \tag{5.8.8}$$

for each unit vector $x \in \mathcal{H}$.

Proof. It follows from (5.8.2) by utilizing the functional calculus that

$$c(A^2 + x_0^2 I - 2x_0 A) + lA - lx_0 I + f(x_0)I \le f(A), \tag{5.8.9}$$

which is equivalent to

$$c(\langle A^2x, x\rangle + x_0^2 - 2x_0\langle Ax, x\rangle) + l\langle Ax, x\rangle - lx_0 + f(x_0) \le \langle f(A)x, x\rangle, \tag{5.8.10}$$

for each unit vector $x \in \mathcal{H}$.

Now, by applying (5.8.10) for $x_0 = \langle Ax, x\rangle$, we deduce the desired inequality (5.8.8). □

Remark 5.8.5. Note that if $A \ge 0$, then $\langle A^2x, x\rangle - \langle Ax, x\rangle^2 \ge 0$. Therefore, we have

$$f(\langle Ax, x\rangle) \le \langle f(A)x, x\rangle - c(\langle A^2x, x\rangle - \langle Ax, x\rangle^2) \le \langle f(A)x, x\rangle.$$

Remark 5.8.6 (Applications for Hölder–McCarthy's inequality).

(i) Consider the function $f : (1, \infty) \to \mathbb{R}, f(x) = x^r$ with $r \ge 2$. It can be easily verified that this function is strongly convex with modulus $c = \frac{r^2-r}{2}$. Based on this fact, from Theorem 5.8.4 we obtain

$$\langle Ax, x\rangle^r \le \langle A^r x, x\rangle - \frac{(r^2 - r)}{2}(\langle A^2x, x\rangle - \langle Ax, x\rangle^2), \tag{5.8.11}$$

for each positive operator A with $Sp(A) \subset (1, \infty)$ and a unit vector $x \in \mathcal{H}$. It is obvious that the inequality (5.8.11) is a refinement of Theorem 5.8.3(i).

(ii) It is readily checked that the function $f : (0, 1) \to \mathbb{R}, f(x) = -x^r$ with $0 < r < 1$ is a strongly convex function with modulus $c = \frac{r-r^2}{2}$. Similar to the above, by using Theorem 5.8.4 we get

$$\langle A^r x, x\rangle \le \langle Ax, x\rangle^r + \frac{(r - r^2)}{2}(\langle Ax, x\rangle^2 - \langle A^2x, x\rangle), \tag{5.8.12}$$

for each positive operator A with $Sp(A) \subset (0, 1)$ and a unit vector $x \in \mathcal{H}$. Apparently, inequality (5.8.12) is a refinement of Theorem 5.8.3(ii).

Now, we draw special attention to the case f^{ν}, $(0 \le \nu \le 1)$. It is strongly convex for which further refinement is possible.

Theorem 5.8.5 ([214]). *Let $f : J \to \mathbb{R}$ be nonnegative and strongly convex with modulus c. If f^{ν} is strongly convex with modulus c' for $\nu \in [0,1]$, then*

$$\begin{aligned} f(\langle Ax,x\rangle) &\le f^{1-\nu}(\langle Ax,x\rangle)f^{\nu}(\langle Ax,x\rangle) + c' f^{1-\nu}(\langle Ax,x\rangle)(\langle A^2x,x\rangle - \langle Ax,x\rangle^2) \\ &\le f^{1-\nu}(\langle Ax,x\rangle)\langle f^{\nu}(A)x,x\rangle \le f^{1-\nu}(\langle Ax,x\rangle)\langle f(A)x,x\rangle^{\nu} \\ &\le (1-\nu)f(\langle Ax,x\rangle) + \nu\langle f(A)x,x\rangle \\ &\le \langle f(A)x,x\rangle - c(1-\nu)(\langle A^2x,x\rangle - \langle Ax,x\rangle^2) \le \langle f(A)x,x\rangle, \end{aligned} \tag{5.8.13}$$

for any positive operator A with $Sp(A) \subset J$ and unit vector $x \in \mathcal{H}$.

Proof. As we have assumed above f^{ν} is strongly convex, so (5.8.8) gives

$$f^{\nu}(\langle Ax,x\rangle) \le \langle f^{\nu}(A)x,x\rangle - c'(\langle A^2x,x\rangle - \langle Ax,x\rangle^2).$$

Multiplying both sides by $f^{1-\nu}(\langle Ax,x\rangle)$, we infer that

$$\begin{aligned} f(\langle Ax,x\rangle) &= f^{1-\nu}(\langle Ax,x\rangle)f^{\nu}(\langle Ax,x\rangle) \\ &\le f^{1-\nu}(\langle Ax,x\rangle)\langle f^{\nu}(A)x,x\rangle - c' f^{1-\nu}(\langle Ax,x\rangle)(\langle A^2x,x\rangle - \langle Ax,x\rangle^2). \end{aligned}$$

Rearranging the terms, we obtain

$$\begin{aligned} f(\langle Ax,x\rangle) &\le f^{1-\nu}(\langle Ax,x\rangle)f^{\nu}(\langle Ax,x\rangle) + c' f^{1-\nu}(\langle Ax,x\rangle)(\langle A^2x,x\rangle - \langle Ax,x\rangle^2) \\ &\le f^{1-\nu}(\langle Ax,x\rangle)\langle f^{\nu}(A)x,x\rangle. \end{aligned} \tag{5.8.14}$$

On the other hand, by using Hölder–McCarthy inequality we have

$$\langle f^{\nu}(A)x,x\rangle \le \langle f(A)x,x\rangle^{\nu}.$$

Multiplying both sides by $f^{1-\nu}(\langle Ax,x\rangle)$, we obtain

$$\begin{aligned} f^{1-\nu}(\langle Ax,x\rangle)\langle f^{\nu}(A)x,x\rangle &\le f^{1-\nu}(\langle Ax,x\rangle)\langle f(A)x,x\rangle^{\nu} \\ &\le (1-\nu)f(\langle Ax,x\rangle) + \nu\langle f(A)x,x\rangle \quad \text{(by Young's inequality)} \\ &\le \langle f(A)x,x\rangle - c(1-\nu)(\langle A^2x,x\rangle - \langle Ax,x\rangle^2) \quad \text{(by (5.8.8))}. \end{aligned} \tag{5.8.15}$$

Combining (5.8.14) and (5.8.15) yields the desired result (5.8.13). □

By (5.8.3) and in a similar manner to the proof of Theorem 5.8.4, we have the following additive reverse.

Theorem 5.8.6 ([214]). *Let $f : J \to \mathbb{R}$ be strongly convex with modulus c and differentiable on J^o whose derivative f' is continuous on J^o. If A is a self-adjoint operator on the Hilbert space $\mathcal{H}$ with $Sp(A) \subset J^o$, then we have*

$$\langle A^2x, x\rangle - \langle Ax, x\rangle^2 \le \frac{1}{2c}\big(\langle f'(A)Ax, x\rangle - \langle Ax, x\rangle\langle f'(A)x, x\rangle\big),$$

for each unit vector $x \in \mathcal{H}$.

By replacing $c(x-y)^2$ with a nonnegative real valued function $F(x-y)$, we can define F**-strongly convex function** as follows:

$$f\big(vx + (1-v)y\big) \le vf(x) + (1-v)f(y) - v(1-v)F(x-y), \tag{5.8.16}$$

for each $v \in [0,1]$ and $x, y \in J$. (This approach has been investigated by M. Adamek in [2].)

We should note that, if F is F-*strongly affine* (i. e., "=" instead of "≤" in (5.8.16)) then the function f is F-strongly convex if and only if $g = f - F$ is convex (see [2, Lemma 4]).

From (5.8.16), we infer that

$$f\big(v(x-y)+y\big) - f(y) + v(1-v)F(x-y) \le v\big(f(x) - f(y)\big).$$

By dividing both sides by v, we obtain

$$\frac{f(v(x-y)+y) - f(y)}{v} + (1-v)F(x-y) \le f(x) - f(y).$$

Notice that if f is differentiable, then by letting $v \to 0$ we find that

$$f'(y)(x-y) + F(x-y) + f(y) \le f(x), \tag{5.8.17}$$

for all $x, y \in J$ and $v \in [0,1]$.

In a similar manner to the proof of Theorem 5.8.4, if F is a continuous function it follows from (5.8.17) that

$$f(\langle Ax, x\rangle) \le \langle f(A)x, x\rangle - \langle F(A - \langle Ax, x\rangle)x, x\rangle, \tag{5.8.18}$$

for any self-adjoint operator A and $x \in \mathcal{H}$, with $\|x\| = 1$. The following theorem is a generalization of (5.8.18). The idea of the proof, given below for completion, is similar to that in [155, Lemma 2.3].

Theorem 5.8.7 ([214]). *Let $f : J \to \mathbb{R}$ be an F-strongly convex and differentiable function on J^o and let $F : J \to [0, \infty)$ be a continuous function. If A is a self-adjoint operator on the Hilbert space $\mathcal{H}$ with $Sp(A) \subset J^o$ and $f(0) \le 0$, then we have*

$$f(\langle Ax, x\rangle) \le \langle f(A)x, x\rangle - \left\langle F\left(A - \frac{1}{\|x\|^2}\langle Ax, x\rangle\right)x, x\right\rangle + (\|x\|^4 - \|x\|^2)F\left(\frac{1}{\|x\|^2}\langle Ax, x\rangle\right),$$

for each $x \in \mathcal{H}$ and $\|x\| \le 1$.

Proof. Let $y = \frac{x}{\|x\|}$ so that $\|y\| = 1$, where

$$
\begin{aligned}
&f(\langle Ax, x\rangle) \\
&\quad = f(\|x\|^2\langle Ay, y\rangle + (1 - \|x\|^2)0) \\
&\quad \le \|x\|^2 f(\langle Ay, y\rangle) + (1 - \|x\|^2)f(0) - \|x\|^2(1 - \|x\|^2)F(\langle Ay, y\rangle - 0) \quad \text{(by (5.8.16))} \\
&\quad \le \|x\|^2 f(\langle Ay, y\rangle) + (\|x\|^4 - \|x\|^2)F(\langle Ay, y\rangle) \quad \text{(since } f(0) \le 0) \\
&\quad \le \|x\|^2(\langle f(A)y, y\rangle - \langle F(A - \langle Ay, y\rangle)y, y\rangle) + (\|x\|^4 - \|x\|^2)F(\langle Ay, y\rangle) \quad \text{(by (5.8.18))} \\
&\quad = \|x\|^2\left(\frac{1}{\|x\|^2}\langle f(A)x, x\rangle - \frac{1}{\|x\|^2}\left\langle F\left(A - \frac{1}{\|x\|^2}\langle Ax, x\rangle\right)x, x\right\rangle\right) \\
&\qquad + (\|x\|^4 - \|x\|^2)F\left(\frac{1}{\|x\|^2}\langle Ax, x\rangle\right) \\
&\quad = \langle f(A)x, x\rangle - \left\langle F\left(A - \frac{1}{\|x\|^2}\langle Ax, x\rangle\right)x, x\right\rangle + (\|x\|^4 - \|x\|^2)F\left(\frac{1}{\|x\|^2}\langle Ax, x\rangle\right).
\end{aligned}
$$

This completes the proof. □

Remark 5.8.7. By taking into account that $F(\cdot)$ is a nonnegative function, we infer that

$$
\begin{aligned}
f(\langle Ax, x\rangle) &\le \langle f(A)x, x\rangle - \left\langle F\left(A - \frac{1}{\|x\|^2}\langle Ax, x\rangle\right)x, x\right\rangle + (\|x\|^4 - \|x\|^2)F\left(\frac{1}{\|x\|^2}\langle Ax, x\rangle\right) \\
&\le \langle f(A)x, x\rangle.
\end{aligned}
$$

6 Reverses for classical inequalities

As we have seen in Chapter 2, and it is known that we have the inequality,

$$(1-v)+vt \geq t^v \tag{6.0.1}$$

for $t > 0$ and $0 \leq v \leq 1$. This inequality states that the weighted arithmetic mean is always greater than or equal to the geometric mean.

For example, multiplying Kantorovich constant $K(t) = \frac{(t+1)^2}{4t} \geq 1$ to the right-hand side in the above inequality, we obtain the *reverse* of the arithmetic–geometric mean inequality, namely we have

$$(1-v)+vt \leq K(t)t^v \tag{6.0.2}$$

for $t > 0$ and $0 \leq v \leq 1$. We easily prove the inequality (6.0.2) in the following. Set the function $f_v(t) = K(t) - \frac{(1-v)+vt}{t^v}$. Then we calculate $f_v'(t) = \frac{(t-1)g_v(t)}{4t^{v+1}}$, where $g_v(t) = t^v + t^{v-1} - 4v(1-v)$. Since $-4v(1-v) \geq -1$ for $0 \leq v \leq 1$, if $t \geq 1$, then $t^v \geq 1$ and $t^{v-1} \geq 0$ and also if $0 < t \leq 1$, then $t^v \geq 0$ and $t^{v-1} \geq 1$, we have $g_v(t) \geq 0$ for $t > 0$ and $0 \leq v \leq 1$. Thus we have $f_v'(t) \leq 0$ for $0 < t \leq 1$ and $f_v'(t) \geq 0$ for $t \geq 1$ which imply $f_v(t) \geq f_v(1) = 0$.

From (6.0.1) and (6.0.2), for $a, b > 0$ we thus have

$$a^{1-v}b^v \leq (1-v)a + vb \leq \frac{(a+b)^2}{4ab}a^{1-v}b^v,$$

by putting $t = b/a$. The first inequality is the arithmetic–geometric mean inequality and the second one is its reverse with Kantorovich constant. There are several constants in the right-hand side. They have been studied in Chapter 2.

In this chapter, we treat with such *reverses* for classical known inequalities.

6.1 Reverses for complemented Golden–Thompson inequalities

For positive semidefinite matrices A and B, the **weak log-majorization** $A \prec_{w\log} B$ means that

$$\prod_{j=1}^{k} \lambda_j(A) \leq \prod_{j=1}^{k} \lambda_j(B), \quad k = 1, 2, \ldots, n,$$

where $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$ are the eigenvalues of A arranged with decreasing order. If the equality holds when $k = n$, then we call it the **log-majorization** $A \prec_{\log} B$. It is known that the weak log-majorization $A \prec_{w\log} B$ implies $\|A\|_u \leq \|B\|_u$ for any **unitarily invariant norm** $\|\cdot\|_u$, that is, $\|UAV\|_u = \|A\|_u$ for any matrix A and all unitary matrices U, V. See [151, 20] for the details of majorization.

https://doi.org/10.1515/9783110643473-006

In the paper [223], Specht obtained the following inequality for $x_1 \geq \cdots \geq x_n > 0$ with $h = x_1/x_n$:

$$\frac{x_1 + \cdots + x_n}{n} \leq S(h)(x_1 \dots x_n)^{\frac{1}{n}},$$

which is a ratio reverse inequality of the arithmetic–geometric mean inequality.

The **Golden–Thompson inequality** [102, 224] is given by $\operatorname{Tr} e^{H+K} \leq \operatorname{Tr} e^H e^K$ for arbitrary Hermitian matrices H and K. In [139, 225], A. Lenard and C. J. Thompson proved the generalized result as

$$\|e^{H+K}\|_u \leq \|e^{H/2}e^K e^{H/2}\|_u \tag{6.1.1}$$

with a unitarily invariant norm. This inequality has been complemented in [114, 7]. T. Ando and F. Hiai in [7] proved that the **complemented Golden–Thompson inequality** for every unitarily invariant norm $\|\cdot\|_u$ and $p > 0$,

$$\|(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}}\|_u \leq \|e^{(1-v)H+vK}\|_u. \tag{6.1.2}$$

Y. Seo in [219] found some upper bounds on $\|e^{(1-v)H+vK}\|_u$ by multiplying constants to $\|(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}}\|_u$, which give reverses of the complemented Golden–Thompson inequality (6.1.2). In this section, we show another reverse of complemented Golden–Thompson inequality, which improve Seo's results. In fact, the general **sandwich condition** $sA \leq B \leq tA$ for positive definite matrices is the key point on results in this section. Also, the so-called **Olson order** $\preceq_{\text{ols}}$ is used. For positive operators A and B, $A \preceq_{\text{ols}} B$ if and only if $A^r \leq B^r$ for every $r \geq 1$ [198].

To study the Golden–Thompson inequality, T. Ando and F. Hiai in [7] developed the following relation on log-majorization:

$$A^r \sharp_v B^r \prec_{\log} (A\sharp_v B)^r, \quad (r \geq 1),$$

which is equivalent to

$$(A^p \sharp_v B^p)^{\frac{1}{p}} \prec_{\log} (A^q \sharp_v B^q)^{\frac{1}{q}}, \quad (0 < q \leq p).$$

By the following lemmas, we obtain a new reverse of these inequalities in terms of eigenvalue inequalities.

Lemma 6.1.1. *For $A, B > 0$ satisfying $sA \leq B \leq tA$ with $0 < s \leq t$, and $v \in [0,1]$, we have*

$$A^r \sharp_v B^r \leq (\max\{S(s), S(t)\})^r (A\sharp_v B)^r, \quad (0 < r \leq 1). \tag{6.1.3}$$

Proof. Let f be an operator monotone function on $[0, \infty)$. Then according to the proof in [99, Theorem 1], we have

$$f(A)\sharp_v f(B) \leq f(c(A\sharp_v B)),$$

where $c = \max\{S(s), S(t)\}$. Putting $f(t) = t^r$, $(0 < r \leq 1)$, we have the inequality in (6.1.3). □

Lemma 6.1.2. *For $A, B > 0$ satisfying $sA \preceq_{\text{ols}} B \preceq_{\text{ols}} tA$ with $0 < s \le t$, and $v \in [0,1]$. Then we have for $k = 1, 2, \ldots, n$,*

$$\lambda_k(A \natural_v B)^r \le \max\{S(s^r), S(t^r)\}\lambda_k(A^r \natural_v B^r), \quad (r \ge 1), \tag{6.1.4}$$

and hence,

$$\lambda_k(A^q \natural_v B^q)^{\frac{1}{q}} \le (\max\{S(s^p), S(t^p)\})^{\frac{1}{p}}\lambda_k(A^p \natural_v B^p)^{\frac{1}{p}}, \quad (0 < q \le p). \tag{6.1.5}$$

Proof. Note that the condition $sA \preceq_{\text{ols}} B \preceq_{\text{ols}} tA$ is equivalent to the condition $s^v A^v \le B^v \le t^v A^v$ for every $v \ge 1$. In particular, we have $sA \le B \le tA$ for $v = 1$. Also, for $r \ge 1$ we have $0 < \frac{1}{r} \le 1$ and by (6.1.3)

$$A^{\frac{1}{r}} \natural_v B^{\frac{1}{r}} \le (\max\{S(s), S(t)\})^{\frac{1}{r}}(A \natural_v B)^{\frac{1}{r}}. \tag{6.1.6}$$

On the other hand, from the condition $s^v A^v \le B^v \le t^v A^v$ for every $v \ge 1$ and letting $v = r$, we have $s^r A^r \le B^r \le t^r A^r$. Now if we take $X = A^r$, $Y = B^r$, $w = s^r$ and $z = t^r$, then

$$wX \le Y \le zX. \tag{6.1.7}$$

Applying (6.1.6) under the condition (6.1.7), we have

$$X^{\frac{1}{r}} \natural_v Y^{\frac{1}{r}} \le (\max\{S(w), S(z)\})^{\frac{1}{r}}(X \natural_v Y)^{\frac{1}{r}},$$

which is equivalent to

$$A \natural_v B \le (\max\{S(s^r), S(t^r)\})^{\frac{1}{r}}(A^r \natural_v B^r)^{\frac{1}{r}}.$$

Thus we have

$$\lambda_k(A \natural_v B) \le (\max\{S(s^r), S(t^r)\})^{\frac{1}{r}}\lambda_k(A^r \natural_v B^r)^{\frac{1}{r}}.$$

By taking rth power on both sides and using the spectral mapping theorem, we get the inequality (6.1.4). Note that from the minimax characterization of eigenvalues of a Hermitian matrix [20, Corollary III.1.2] it follows immediately that $A \le B$ implies $\lambda_k(A) \le \lambda_k(B)$ for each k. Similarly, since $p/q \ge 1$, from inequality (6.1.4), we have

$$\lambda_k(A \natural_v B)^{\frac{p}{q}} \le \max\{S(s^{\frac{p}{q}}), S(t^{\frac{p}{q}})\}\lambda_k(A^{\frac{p}{q}} \natural_v B^{\frac{p}{q}}). \tag{6.1.8}$$

Replacing A and B by A^q and B^q in (6.1.8), and using the sandwich condition $s^q A^q \le B^q \le t^q A^q$, we have

$$\lambda_k(A^q \natural_v B^q)^{\frac{p}{q}} \le \max\{S(s^p), S(t^p)\}\lambda_k(A^p \natural_v B^p).$$

This completes the proof. □

We have the following corollary for the condition $mI \le A, B \le MI$ by the use of Lemma 6.1.1 and Lemma 6.1.2.

Corollary 6.1.1. *For $A, B > 0$ satisfying $mI \le A, B \le MI$ with $0 < m \le M$, $h = M/m$, $v \in [0,1]$ and $k = 1, 2, \ldots, n$, we have*

$$A^r \sharp_v B^r \le S(h)^r (A \sharp_v B)^r, \quad (0 < r \le 1), \tag{6.1.9}$$

and hence

$$\lambda_k (A \sharp_v B)^r \le S(h^r) \lambda_k (A^r \sharp_v B^r), \quad (r \ge 1), \tag{6.1.10}$$

$$\lambda_k (A^q \sharp_v B^q)^{\frac{1}{q}} \le S(h^p)^{\frac{1}{p}} \lambda_k (A^p \sharp_v B^p)^{\frac{1}{p}}, \quad (0 < q \le p). \tag{6.1.11}$$

See [100] for the proof of Corollary 6.1.1. In the sequel, we show a reverse of the complemented Golden–Thompson inequality (6.1.2).

Theorem 6.1.1 ([100]). *Let H and K be Hermitian matrices such that $e^s e^H \preceq_{\mathrm{ols}} e^K \preceq_{\mathrm{ols}} e^t e^H$ with $s \le t$, and let $v \in [0,1]$. Then we have for all $p > 0$ and $k = 1, 2, \ldots, n$,*

$$\lambda_k (e^{(1-v)H + vK}) \le (\max\{S(e^{sp}), S(e^{tp})\})^{\frac{1}{p}} \lambda_k (e^{pH} \sharp_v e^{pK})^{\frac{1}{p}}.$$

Proof. Replacing A and B by e^H and e^K in the inequality (6.1.5) of Lemma 6.1.2, we can write

$$\lambda_k (e^{qH} \sharp_v e^{qK})^{\frac{1}{q}} \le (\max\{S(e^{sp}), S(e^{tp})\})^{\frac{1}{p}} \lambda_k (e^{pH} \sharp_v e^{pK})^{\frac{1}{p}}, \quad (0 < q \le p).$$

By [114, Lemma 3.3], we have

$$e^{(1-v)H + vK} = \lim_{q \to 0} (e^{qH} \sharp_v e^{qK})^{\frac{1}{q}},$$

and hence we have for each $p > 0$,

$$\lambda_k (e^{(1-v)H + vK}) \le (\max\{S(e^{sp}), S(e^{tp})\})^{\frac{1}{p}} \lambda_k (e^{pH} \sharp_v e^{pK})^{\frac{1}{p}}. \qquad \square$$

Note that eigenvalue inequalities immediately imply log-majorization and unitarily invariant norm inequalities.

Corollary 6.1.2. *Let H and K be Hermitian matrices such that $e^s e^H \preceq_{\mathrm{ols}} e^K \preceq_{\mathrm{ols}} e^t e^H$ with $s \le t$, and let $v \in [0,1]$. Then we have for every unitarily invariant norm $\|\cdot\|_u$ and all $p > 0$,*

$$\|e^{(1-v)H + vK}\|_u \le (\max\{S(e^{sp}), S(e^{tp})\})^{\frac{1}{p}} \|(e^{pH} \sharp_v e^{pK})^{\frac{1}{p}}\|_u, \tag{6.1.12}$$

and the right-hand side of (6.1.12) converges to the left-hand side as $p \downarrow 0$. In particular, we have

$$\|e^{H+K}\|_u \le \max\{S(e^{2s}), S(e^{2t})\} \|(e^{2H} \sharp e^{2K})\|_u.$$

It is remarkable that the inequality (6.1.12) does not give a better upper bound of $\|e^{(1-v)H+vK}\|_u$ for the case $p = 1$ compared to the original inequality (6.1.1). However, the inequality (6.1.12) holds for any $p > 0$ and we take the limit $p \to 0$, the right-hand side of (6.1.12) can be close to the left-hand side of (6.1.12) as one hope. This means the inequality (6.1.12) gives a better upper bound of $\|e^{(1-v)H+vK}\|_u$ for a general case of $p > 0$. It may be interesting problem to find the constant $c \leq 1$ such that $\|e^{H+K}\|_u \leq c\|e^{H/2}e^K e^{H/2}\|_u$ as the special case of $p = 1$.

Corollary 6.1.3. *Let H and K be Hermitian matrices such that $mI \leq H, K \leq MI$ with $m \leq M$, and let $v \in [0,1]$. Then we have for all $p > 0$,*

$$\lambda_k(e^{(1-v)H+vK}) \leq S^{1/p}(e^{(M-m)p})\lambda_k(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}}, \quad (k = 1, 2, \ldots, n).$$

We thus have for every unitarily invariant norm $\|\cdot\|_u$,

$$\|e^{(1-v)H+vK}\|_u \leq S^{1/p}(e^{(M-m)p})\|(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}}\|_u,$$

and the right-hand side of these inequalities converges to the left-hand side as $p \downarrow 0$.

See [100] for the proof of Corollary 6.1.3 and [219, Theorem 3.3–Theorem 3.4] as references for the readers. A well-known matrix version of the **Kantorovich inequality** [153] asserts that if A and U are two matrices such that $0 < mI \leq A \leq MI$ and $UU^* = I$, then

$$UA^{-1}U^* \leq \frac{(m+M)^2}{4mM}(UAU^*)^{-1}. \tag{6.1.13}$$

Now as a result of the following statement with the generalized Kantorovich constant given in (2.0.8), we have an another reverse of the complemented Golden–Thompson inequality which refines corresponding inequality in [219] by Y. Seo.

Proposition 6.1.1. *Let H and K be Hermitian matrices such that $e^s e^H \preceq_{\mathrm{ols}} e^K \preceq_{\mathrm{ols}} e^t e^H$ for some scalars $s \leq t$, and let $v \in [0,1]$. Then*

$$\lambda_k(e^{(1-v)H+vK}) \leq K^{-1/p}(e^{p(t-s)}, v)\lambda_k(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}}, \quad p > 0, \tag{6.1.14}$$

where $K(\cdot,\cdot)$ is the generalized Kantorovich constant defined as (2.0.8).

Theorem 6.1.2 ([100]). *Let H and K be Hermitian matrices such that $mI \leq K, H \leq MI$ with $m \leq M$ and let $v \in [0,1]$. Then we have for every $p > 0$,*

$$\lambda_k(e^{(1-v)H+vK}) \leq K^{-1/p}(e^{2p(M-m)}, v)\lambda_k(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}}, \quad (k = 1, 2, \ldots, n),$$

and the right-hand side of this inequality converges to the left-hand side as $p \downarrow 0$. In particular, we have

$$\lambda_k(e^{H+K}) \leq \frac{e^{2M} + e^{2m}}{2e^M e^m}\lambda_k(e^{2H}\sharp e^{2K}), \quad (k = 1, 2, \ldots, n).$$

Proof. Since $mI \le K, H \le MI$ implies $e^{m-M}e^H \preceq_{\mathrm{ols}} e^K \preceq_{\mathrm{ols}} e^{M-m}e^H$, desired inequalities are obtained by letting $s = m - M$ and $t = M - m$ in Proposition 6.1.1. For the convergence, we know that $\frac{2w^{1/4}}{w^{1/2}+1} \le K(w, v) \le 1$, for every $v \in [0, 1]$ by (iv) and (v) in Lemma 2.0.2. We have for every $p > 0$, $1 \le K^{-1/p}(w^p, v) \le (\frac{2w^{p/4}}{w^{p/2}+1})^{-1/p}$. Simple calculations show that $\lim_{p\to 0} -\frac{1}{p}\log(\frac{2w^{p/4}}{w^{p/2}+1}) = \lim_{p\to 0} \frac{(w^{p/2}-1)\log w}{4w^{p/2}+1} = 0$, and then $\lim_{p\to 0}(\frac{2w^{p/4}}{w^{p/2}+1})^{-1/p} = 1$. From the sandwich principle with $w = e^{2(M-m)}$, we have $\lim_{p\to 0} K(e^{2p(M-m)}, v)^{-1/p} = 1$. □

Remark 6.1.1. Under the assumptions of Theorem 6.1.2, Y. Seo in [219, Theorem 4.2] proved that

$$\|e^{(1-v)H+vK}\|_u \le K^{-v/p}(e^{(M-m)}, p)K^{-1/p}(e^{2p(M-m)}, v)\|(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}}\|_u, \quad (0 < p \le 1),$$

and

$$\|e^{(1-v)H+vK}\|_u \le K^{-1/p}(e^{2p(M-m)}, v)\|(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}}\|_u, \quad (p \ge 1).$$

However, the inequality (6.1.14) shows that the sharper constant is $K^{-\frac{1}{p}}(e^{2p(M-m)}, v)$ for all $p > 0$. Since for $0 < p \le 1$, $K^{-v/p}(e^{(M-m)}, p) \ge 1$, and hence

$$K^{-1/p}(e^{2p(M-m)}, v) \le K^{-v/p}(e^{(M-m)}, p)K^{-1/p}(e^{2p(M-m)}, v).$$

It has been shown in [99] that if $f : [0, \infty) \longrightarrow [0, \infty)$ is an operator monotone function and $0 < mI \le A \le B \le MI \le I$ with $h = \frac{M}{m}$, then for all $v \in [0, 1]$

$$f(A)\sharp_v f(B) \le \exp(v(1-v)(1-h)^2)f(A\sharp_v B). \tag{6.1.15}$$

This new ratio has been introduced by S. Furuichi and N. Minculete in [77] (see Theorem 2.3.1), which is different from Specht ratio and Kantorovich constant. By applying (6.1.15) for $f(t) = t^r$, $0 < r \le 1$, we have the following results similar to Lemma 6.1.1 and Lemma 6.1.2.

Lemma 6.1.3. *For $A, B > 0$ satisfying $0 < mI \le A \le B \le MI \le I$ with $h = M/m$, and $v \in [0, 1]$, we have*

$$A^r \sharp_v B^r \le \exp(rv(1-v)(1-h)^2)(A\sharp_v B)^r, \quad (0 < r \le 1).$$

Lemma 6.1.4. *For $A, B > 0$ satisfying $0 < mI \preceq_{\mathrm{ols}} A \preceq_{\mathrm{ols}} B \preceq_{\mathrm{ols}} MI \preceq_{\mathrm{ols}} I$ with $h = M/m$, $v \in [0, 1]$, and for all $k = 1, 2, \dots, n$, we have*

$$\lambda_k(A\sharp_v B)^r \le \exp(v(1-v)(1-h^r)^2)\lambda_k(A^r\sharp_v B^r), \quad (r \ge 1),$$
$$\lambda_k(A^q\sharp_v B^q)^{\frac{1}{q}} \le \exp\left(\frac{1}{p}v(1-v)(1-h^p)^2\right)\lambda_k(A^p\sharp_v B^p)^{\frac{1}{p}}, \quad (0 < q \le p). \tag{6.1.16}$$

Theorem 6.1.3 ([100]). *Let H and K be Hermitian matrices such that $e^m I \preceq_{\mathrm{ols}} e^H \preceq_{\mathrm{ols}} e^K \preceq_{\mathrm{ols}} e^M I \preceq_{\mathrm{ols}} I$ with $m \leq M$, and let $v \in [0,1]$. Then we have for all $p > 0$ and $k = 1, 2, \ldots, n$,*

$$\lambda_k(e^{(1-v)H+vK}) \leq \exp\left(\frac{1}{p}v(1-v)(1-e^{p(M-m)})^2\right)\lambda_k(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}},$$

and we thus have, for every unitarily invariant norm $\|\cdot\|_u$,

$$\|e^{(1-v)H+vK}\|_u \leq \exp\left(\frac{1}{p}v(1-v)(1-e^{p(M-m)})^2\right)\|(e^{pH}\sharp_v e^{pK})^{\frac{1}{p}}\|_u.$$

Proof. The proof is similar to that of Theorem 6.1.1, by replacing A and B with e^H and e^K, and $h = e^{M-m}$ in the inequality (6.1.16). □

6.2 Reverses for operator Aczél inequality

In the paper [1], J. Aczél proved that if $a_i, b_i (1 \leq i \leq n)$ are positive real numbers such that $a_1^2 - \sum_{i=2}^n a_i^2 > 0$ and $b_1^2 - \sum_{i=2}^n b_i^2 > 0$, then

$$\left(a_1^2 - \sum_{i=2}^n a_i^2\right)\left(b_1^2 - \sum_{i=2}^n b_i^2\right) \leq \left(a_1 b_1 - \sum_{i=2}^n a_i b_i\right)^2.$$

Aczél inequality has important applications in the theory of functional equations in non-Euclidean geometry [1] and references therein. Some considerable attentions have been given to this inequality involving its generalizations, variations and applications. T. Popoviciu [204] first presented an exponential extension for Aczél inequality as follows.

Theorem 6.2.1 ([204]). *Let $p > 0, q > 0$ with $\frac{1}{p} + \frac{1}{q} = 1$, $a_1^p - \sum_{i=2}^n a_i^p > 0$, and $b_1^q - \sum_{i=2}^n b_i^q > 0$. Then we have*

$$\left(a_1^p - \sum_{i=2}^n a_i^p\right)^{\frac{1}{p}}\left(b_1^q - \sum_{i=2}^n b_i^q\right)^{\frac{1}{q}} \leq a_1 b_1 - \sum_{i=2}^n a_i b_i.$$

The above inequality is often called **Popoviciu inequality**. S. S. Dragomir [39] established a variant of Aczél inequality in an inner product space.

Theorem 6.2.2 ([39]). *Let a, b be real numbers and x, y be vectors of an inner product space such that $a^2 - \|x\|^2 > 0$ or $b^2 - \|y\|^2 > 0$. Then*

$$(a^2 - \|x\|^2)(b^2 - \|y\|^2) \leq (ab - \mathrm{Re}\langle x, y\rangle)^2. \tag{6.2.1}$$

M. S. Moslehian in [184] proved an **operator Aczél inequality** involving the weighted geometric operator mean.

Theorem 6.2.3 ([184]). *Let* $g : (0,\infty) \to [0,\infty)$ *be an operator decreasing and operator concave function,* $\frac{1}{p} + \frac{1}{q} = 1$, $p, q > 1$. *For* $A, B > 0$ *and all* $\xi \in \mathcal{H}$, *we then have*

$$g(A^p)\sharp_{\frac{1}{q}} g(B^q) \le g(A^p \sharp_{\frac{1}{q}} B^q), \tag{6.2.2}$$

$$\langle g(A^p)\xi, \xi\rangle^{\frac{1}{p}} \langle g(B^q)\xi, \xi\rangle^{\frac{1}{q}} \le \langle g(A^p \sharp_{\frac{1}{q}} B^q)\xi, \xi\rangle. \tag{6.2.3}$$

In this section, we present some reverses for operator Aczél inequality (6.2.2) and (6.2.3), by using several reverse Young inequalities. In fact, we show some upper bounds for inequalities in Theorem 6.2.3. These results are proved for a nonnegative operator monotone decreasing function g without operator concavity. That is, we use less restrictive conditions on g. Actually, the conditions on g in Theorem 6.2.3 seem to be rigid restrictions because the operator concavity is equivalent to the operator monotone increasingness, if $g > 0$.

The study on Young inequality is interesting and there are several ratio-type and difference-type reverses of this inequality [42, 140]. We have studied the recent advances of Young inequalities in Chapter 2. One of reverse inequalities is given by W. Liao et al. [140] using the Kantorovich constant as follows.

Lemma 6.2.1 ([140, Theorem 3.1]). *For* $A, B > 0$ *satisfying the conditions either* $0 < mI \le A \le m'I \le M'I \le B \le MI$ *or* $0 < mI \le B \le m'I \le M'I \le A \le MI$, *with some constants* m, m', M, M', *we have*

$$(1-v)A + vB \le K^R(h)(A\sharp_v B), \tag{6.2.4}$$

where $h = \frac{M}{m}$, $v \in [0,1]$, $R = \max\{1-v, v\}$.

In the following, we generalize Lemma 6.2.1 with the more general sandwich condition $0 < sA \le B \le tA$. The proof is similar to that of [226, Theorem 2.1].

Lemma 6.2.2. *Let* $0 < sA \le B \le tA$ *with* $0 < s \le t$ *and let* $v \in [0,1]$. *Then we have*

$$(1-v)A + vB \le \max\{K^R(s), K^R(t)\}(A\sharp_v B), \tag{6.2.5}$$

where $R = \max\{v, 1-v\}$.

Proof. From [140, Corollary 2.2], for $x > 0$ and $v \in [0,1]$, we have

$$(1-v) + vx \le K^R(x)x^v.$$

Thus for every strictly positive operator $0 < sI \le C \le tI$, we have

$$(1-v) + vC \le \max_{s\le x\le t} K^R(x)C^v.$$

Substituting $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$ for C, we get

$$(1-v) + vA^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le \max_{s\le x\le t} K^R(x)(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^v.$$

Multiplying $A^{\frac{1}{2}}$ to the both sides in the above inequality, and using the fact that $\max_{s\le x\le t} K(x) = \max\{K(s), K(t)\}$, the desired inequality is obtained. ☐

Remark 6.2.1. Lemma 6.2.2 is a generalization of Lemma 6.2.1. Since, if $0 < mI \le A \le m'I \le M'I \le B \le MI$ or $0 < mI \le B \le m'I \le M'I \le A \le MI$, then $\frac{m}{M}A \le B \le \frac{M}{m}A$. Now by letting $s = \frac{m}{M}$ and $t = \frac{M}{m}$ in Lemma 6.2.2, the inequality (6.2.4) is obtained.

Lemma 6.2.3. *Let $g : (0,\infty) \to [0,\infty)$ be an operator monotone decreasing function. Then, for $A > 0$ and $\lambda \ge 1$,*

$$\frac{1}{\lambda} g(A) \le g(\lambda A).$$

Proof. First, note that since g is analytic on $(0,\infty)$, we may assume that $g(x) > 0$ for all $x > 0$. Otherwise, g is identically zero. Since g is also an operator monotone decreasing on $(0,\infty)$, $f = 1/g$ is nonnegative operator monotone on $(0,\infty)$, and thus operator concave function [20, Theorem V.2.5]. On the other hand, it is known that for every nonnegative concave function f and $\lambda \ge 1$, $f(\lambda x) \le \lambda f(x)$. Therefore, for every $\lambda \ge 1$ we have

$$\left(g(\lambda A)\right)^{-1} \le \lambda \left(g(A)\right)^{-1}.$$

Taking inverse of this inequality, we obtain the result. ☐

Proposition 6.2.1. *Let $g : (0,\infty) \to [0,\infty)$ be an operator monotone decreasing function and $0 < sA \le B \le tA$ with $0 < s \le t$. Then we have for all $v \in [0,1]$,*

$$g(A \sharp_v B) \le \max\{K^R(s), K^R(t)\}\left(g(A)\sharp_v g(B)\right), \tag{6.2.6}$$

where $R = \max\{v, 1-v\}$.

Proof. Since $0 < sA \le B \le tA$, from Lemma 6.2.2 we have

$$(1-v)A + vB \le \lambda(A \sharp_v B),$$

where $\lambda = \max\{K^R(s), K^R(t)\}$. We know that $\lambda \ge 1$. Also, the function g is operator monotone decreasing and so

$$g\left(\lambda(A \sharp_v B)\right) \le g\left((1-v)A + vB\right).$$

We can write

$$\frac{1}{\lambda} g(A \sharp_v B) \le g\left(\lambda(A \sharp_v B)\right) \le g\left((1-v)A + vB\right) \le g(A)\sharp_v g(B),$$

where the first inequality follows from Lemma 6.2.3 and the last inequality follows from [8, Remark 2.6]. ☐

Now, we can give one of **reverses of operator Aczél inequality** (with Kantorovich constant) for Theorem 6.2.3 as follows.

Theorem 6.2.4 ([100]). *Let $g : (0,\infty) \to [0,\infty)$ be an operator monotone decreasing function, $\frac{1}{p} + \frac{1}{q} = 1$, $p, q > 1$ and $0 < sA^p \le B^q \le tA^p$ with $0 < s \le t$. Then we have for all $\xi \in \mathcal{H}$,*

$$g(A^p \sharp_{\frac{1}{q}} B^q) \le \max\{K^R(s), K^R(t)\}(g(A^p) \sharp_{\frac{1}{q}} g(B^q)), \tag{6.2.7}$$

$$\langle g(A^p \sharp_{\frac{1}{q}} B^q)\xi, \xi\rangle \le \max\{K^R(s), K^R(t)\}\langle g(A^p)\xi, \xi\rangle^{\frac{1}{p}} \langle g(B^q)\xi, \xi\rangle^{\frac{1}{q}}, \tag{6.2.8}$$

where $R = \max\{1/p, 1/q\}$.

Proof. Letting $v = 1/q$ and replacing A^p and B^q with A and B in Proposition 6.2.1, we have the inequality (6.2.7). To prove the inequality (6.2.8), first note that under the condition $0 < sA^p \le B^q \le tA^p$ from Lemma 6.2.2 we have

$$A^p \nabla_v B^q \le \max\{K^R(s), K^R(t)\}(A^p \sharp_v B^q).$$

For convenience, set $\lambda = \max\{K^R(s), K^R(t)\}$. Therefore, for the operator monotone decreasing functions g and $v = 1/q$, we have

$$g(\lambda(A^p \sharp_{1/q} B^q)) \le g(A^p \nabla_{1/q} B^q). \tag{6.2.9}$$

We compute

$$\begin{aligned}\langle g(A^p \sharp_{1/q} B^q)\xi, \xi\rangle &\le \lambda\langle g(\lambda(A^p \sharp_{1/q} B^q))\xi, \xi\rangle \le \lambda\langle g(A^p \nabla_{1/q} B^q))\xi, \xi\rangle \\ &\le \lambda\langle g(A^p) \sharp_{1/q} g(B^q)\xi, \xi\rangle \le \lambda\langle g(A^p)\xi, \xi\rangle^{1/p}\langle g(B^q)\xi, \xi\rangle^{1/q},\end{aligned}$$

where the first inequality follows from Lemma 6.2.3 and the second follows from the inequality (6.2.9). For the third inequality, we use log-convexity property of operator monotone decreasing functions [8, Remark 2.6], and in the last inequality we use the fact that for every positive operators A, B and every $\xi \in \mathcal{H}$, $\langle A \sharp_v B\xi, \xi\rangle \le \langle A\xi, \xi\rangle^{1-v}\langle B\xi, \xi\rangle^v$, [29, Lemma 8]. We thus obtained

$$\langle g(A^p \sharp_{1/q} B^q)\xi, \xi\rangle \le \max\{K^R(s), K^R(t)\}\langle g(A^p)\xi, \xi\rangle^{1/p}\langle g(B^q)\xi, \xi\rangle^{1/q}. \qquad \square$$

Corollary 6.2.1. *Let $\frac{1}{p} + \frac{1}{q} = 1$, $p, q \ge 1$, and A, B be commuting positive operators with spectra contained in $(0, 1)$ such that $0 < sA^p \le B^q \le tA^p$ with $0 < s \le t$. Then we have for every unit vector $\xi \in \mathcal{H}$*

$$1 - \|(AB)^{1/2}\xi\|^2 \le \max\{K^R(s), K^R(t)\}(1 - \|A^{p/2}\xi\|^2)^{1/p}(1 - \|B^{q/2}\xi\|^2)^{1/q}, \tag{6.2.10}$$

and consequently

$$1 - \|AB\xi\|^2 \le \max\{K^R(s^2), K^R(t^2)\}(1 - \|A^p\xi\|^2)^{1/p}(1 - \|B^q\xi\|^2)^{1/q},$$

where $R = \max\{1/p, 1/q\}$.

See [123] for the proof of Corollary 6.2.1.

Remark 6.2.2. M. S. Moslehian in [184, Corollary 2.4], showed operator version of Aczél inequity (6.2.1) as follows:

$$(1 - \|A^{\frac{p}{2}}\xi\|^2)^{\frac{1}{p}}(1 - \|B^{\frac{q}{2}}\xi\|^2)^{\frac{1}{q}} \leq 1 - \|(AB)^{\frac{1}{2}}\xi\|^2, \tag{6.2.11}$$

where A and B are commuting positive operators with spectra contained in $(0, 1)$, and $\frac{1}{p} + \frac{1}{q} = 1$ for $p, q \geq 1$. As it has been seen, the inequality (6.2.10) in Corollary 6.2.1 provides an upper bound for the operator Aczél inequality (6.2.11).

Corollary 6.2.2. *Let $g : (0, \infty) \to [0, \infty)$ be an operator monotone decreasing function and A, B be commuting positive operators such that $0 < sA^p \leq B^q \leq tA^p$ with $0 < s \leq t$. Then we have*

$$g(AB) \leq \max\{K^R(s), K^R(t)\} g(A^p)^{\frac{1}{p}} g(B^q)^{\frac{1}{q}},$$

where $R = \max\{1/p, 1/q\}$.

Corollary 6.2.3. *Let $g : (0, \infty) \to [0, \infty)$ be a decreasing function and a_i, b_i be positive numbers such taht $0 < s \leq \frac{(b_i)^q}{(a_i)^p} \leq t$ with constants s, t. Then we have*

$$\sum_{i=1}^{n} g(a_i b_i) \leq \max\{K^R(s), K^R(t)\} \left(\sum_{i=1}^{n} g(a_i^p)\right)^{1/p} \left(\sum_{i=1}^{n} g(b_i^q)\right)^{1/q}, \tag{6.2.12}$$

where $R = \max\{1/p, 1/q\}$.

See [123] for the proof of Corollary 6.2.3. We state some related results. S. S. Dragomir in [42, Theorem 6], gave an another reverse inequality for Young inequality as follows (we treated the improvements of this kind of inequality in subsection 2.5).

Lemma 6.2.4. *For $A, B > 0$ satisfying $0 < sA \leq B \leq tA$ with $0 < s \leq t$, and for all $v \in [0, 1]$, we have*

$$(1 - v)A + vB \leq \exp\left(\frac{1}{2}v(1 - v)\left(\frac{\max\{1, t\}}{\min\{1, s\}} - 1\right)^2\right)(A \sharp_v B). \tag{6.2.13}$$

By using this new ratio, we can express some other operator reverse inequalities. The proofs are similar to that of preceding results so that we omit the proof. See [123] for the proof of Proposition 6.2.2.

Proposition 6.2.2. *Let $g : (0, \infty) \to [0, \infty)$ be an operator monotone decreasing function and $0 < sA \leq B \leq tA$ with $0 < s \leq t$. Then we have for all $v \in [0, 1]$,*

$$g(A \sharp_v B) \leq \exp\left(\frac{1}{2}v(1 - v)\left(\frac{\max\{1, t\}}{\min\{1, s\}} - 1\right)^2\right)(g(A) \sharp_v g(B)).$$

We can give a **reverses of operator Aczél inequality** (with an alternative constant) for Theorem 6.2.3 as follows.

Theorem 6.2.5 ([100]). *Let $g : (0,\infty) \to [0,\infty)$ be an operator monotone decreasing function, $\frac{1}{p} + \frac{1}{q} = 1$, $p, q > 1$ and $0 < sA^p \le B^q \le tA^p$ with $0 < s \le t$. Then we have for all $\xi \in \mathcal{H}$,*

$$g(A^p \sharp_{1/q} B^q) \le \exp\left(\frac{1}{2pq}\left(\frac{\max\{1,t\}}{\min\{1,s\}} - 1\right)^2\right)(g(A^p)\sharp_{1/q} g(B^q)),$$
$$\langle g(A^p \sharp_{1/q} B^q)\xi,\xi\rangle \le \exp\left(\frac{1}{2pq}\left(\frac{\max\{1,t\}}{\min\{1,s\}} - 1\right)^2\right)\langle g(A^p)\xi,\xi\rangle^{\frac{1}{p}}\langle g(B^q)\xi,\xi\rangle^{\frac{1}{q}}.$$

In [79, Theorem B], an another reverse Young's inequality is presented as follows.

Lemma 6.2.5. *Let A and B be positive operators such that $0 < sA \le B \le A$ for a constant $s > 0$ and $v \in [0,1]$. Then*

$$(1-v)A + vB \le M_v(s)(A \sharp_v B),$$

where $M_v(s) = 1 + \frac{v(1-v)(s-1)^2}{2s^{v+1}}$.

Now by using this new constant, the similar reverse Aczél inequalities are obtained. Note that $M_v(s) \ge 1$ for every $v \in [0,1]$. See Section 2.6 and [79] for more properties of $M_v(s)$.

Proposition 6.2.3. *Let $g : (0,\infty) \to [0,\infty)$ be an operator monotone decreasing function, $0 < sA \le B \le A$ with $s > 0$ and $v \in [0,1]$. Then we have*

$$g(A \sharp_v B) \le M_v(s)\big(g(A)\sharp_v g(B)\big).$$

We can give a **reverses of operator Aczél inequality** (with an alternative constant and restricted assumption) for Theorem 6.2.3 as follows.

Theorem 6.2.6 ([100]). *Let $g : (0,\infty) \to [0,\infty)$ be an operator monotone decreasing function, $\frac{1}{p} + \frac{1}{q} = 1$, $p, q > 1$ and $0 < sA^p \le B^q \le A^p$ for a constant $s > 0$. Then we have for all $\xi \in \mathcal{H}$,*

$$g(A^p \sharp_{1/q} B^q) \le M_{1/q}(s)(g(A^p)\sharp_{1/q} g(B^q)),$$
$$\langle g(A^p \sharp_{1/q} B^q)\xi,\xi\rangle \le M_{1/q}(s)\langle g(A^p)\xi,\xi\rangle^{1/p}\langle g(B^q)\xi,\xi\rangle^{1/q}.$$

Remark 6.2.3. We clearly see that the condition $0 < sA \le B \le tA$ for $0 < s \le t$ in Lemma 6.2.4 is more general than the condition $0 < sA \le B \le A$ for $s \le 1$ in Lemma 6.2.5. But under the same condition $0 < sA \le B \le A$, the appeared constant in Lemma 6.2.5 gives a better estimate than ones in Lemma 6.2.4. In fact, we have

$$M_v(s) \le \exp\left(\frac{1}{2}v(1-v)\left(\frac{1}{s} - 1\right)^2\right),$$

for every $v \in [0,1]$ and $0 < s \le 1$, which is given in [71, Proposition 2.10].

In the paper [99], it is shown that if $f : [0,\infty) \to [0,\infty)$ is an operator monotone function and $0 < sA \le B \le tA$ with $0 < s \le t$, then we have for all $v \in [0,1]$,

$$f(A)\sharp_v f(B) \le \max\{S(s), S(t)\} f(A\sharp_v B),$$

where $S(t)$ is the Specht ratio [97, 223]. As a result, we can show for an operator monotone decreasing function $g : (0,\infty) \to [0,\infty)$, $0 < sA \le B \le tA$, and $v \in [0,1]$,

$$g(A\sharp_v B) \le \max\{S(s), S(t)\}\big(g(A)\sharp_v g(B)\big). \tag{6.2.14}$$

Thus we can obtain the alternative **reverse of operator Aczél inequality** with Specht ratio $\max\{S(p), S(q)\}$.

6.3 Reverses for Bellman operator inequality

The following inequality is well known in the literature as the **operator Bellman inequality** [185]

$$\Phi^{1/p}(I - A\nabla_v B) \ge \Phi\big((I-A)^{1/p}\nabla_v (I-B)^{1/p}\big), \tag{6.3.1}$$

for $v \in [0,1]$, $p > 1$, $0 < A, B \le I$ and a normalized positive linear map Φ. This inequality was proved in [185] by M. S. Moslehian et al. as an operator version of the **scalar Bellman inequality** [17]

$$\left(a^p - \sum_{k=1}^{n} a_k^p\right)^{1/p} + \left(b^p - \sum_{k=1}^{n} b_k^p\right)^{1/p} \le \left((a+b)^p - \sum_{k=1}^{n}(a_k + b_k)^p\right)^{1/p}, \tag{6.3.2}$$

for the positive numbers a, a_k, b, b_k satisfying $\sum\limits_{k=1}^{n} a_k^p \le a^p$ and $\sum\limits_{k=1}^{n} b_k^p \le b^p$, where $p > 1$.

The proof of (6.3.1) was based on the operator inequality [185]

$$f\big(\Phi(A\nabla_v B)\big) \ge \Phi\big(f(A)\nabla_v f(B)\big), \tag{6.3.3}$$

valid for the operator concave function $f : J \subset (0,\infty) \to (0,\infty)$, the normalized positive linear map Φ and the positive operators A, B whose spectra are contained in the interval J.

In this section, we prove a more elaborated reverse of (6.3.3), valid for concave functions (not necessarily operator concave). This reverse inequality will be used to find a reversed version of (6.3.1) and (6.3.2). Further, we present a simple approach that can be used to prove the scalar Bellman inequality and its reverse. The new approach will be useful in obtaining several refinements for their inequalities.

In the sequel, we present a general inequality by applying Mond–Pečarić method. We refer the reader to [97] as a comprehensive reference of this method.

The following notation will be used in Theorem 6.3.1, for the positive numbers m, M and the function $f : [m, M] \to \mathbb{R}$:

$$a_f = \frac{f(M) - f(m)}{M - m} \quad \text{and} \quad b_f = \frac{Mf(m) - mf(M)}{M - m}.$$

Theorem 6.3.1 ([216]). *Let Φ be a normalized positive linear map on $\mathbb{B}(\mathcal{H})$, $A, B \in \mathbb{B}(\mathcal{H})$ be two positive operators such that $mI \le A, B \le MI$ with $0 < m < M$. If $f, g : [m, M] \to [0, \infty)$ are continuous functions such that f is concave, then for a given $\alpha > 0$,*

$$\alpha g(\Phi(A\nabla_v B)) + \beta I \le \Phi(f(A)\nabla_v f(B)), \tag{6.3.4}$$

where $\beta = \min_{t\in[m,M]} \{a_f t + b_f - \alpha g(t)\}$.

The reverse inequality of (6.3.4) holds when f is a convex function.

Proof. According to the assumptions, we have $f(t) \ge a_f t + b_f$ for any $t \in [m, M]$. A standard functional calculus argument implies $f(A) \ge a_f A + b_f I$ and $f(B) \ge a_f B + b_f I$. Consequently, we infer for any $v \in [0, 1]$,

$$(1 - v)f(A) \ge (1 - v)a_f A + (1 - v)b_f I \quad \text{and} \quad vf(B) \ge va_f B + vb_f I, \tag{6.3.5}$$

and we then have

$$f(A)\nabla_v f(B) \ge a_f(A\nabla_v B) + b_f I. \tag{6.3.6}$$

It follows from the linearity and the normality of Φ that

$$\Phi(f(A)\nabla_v f(B)) \ge a_f\Phi(A\nabla_v B) + b_f I.$$

Therefore, we have

$$\begin{aligned}\Phi(f(A)\nabla_v f(B)) - \alpha g(\Phi(A\nabla_v B)) &\ge a_f\Phi(A\nabla_v B) + b_f I - \alpha g(\Phi(A\nabla_v B)) \\ &\ge \min_{t\in[m,M]} \{a_f t + b_f - \alpha g(t)\}I\end{aligned}$$

which implies the desired inequality (6.3.4). □

A reverse for operator Bellman inequality (6.3.1) is obtained by taking $f(t) = g(t) = (1 - t)^{1/p}$ on $(0, 1)$ with $p > 1$ in Theorem 6.3.1 in the following.

Corollary 6.3.1 (Reverse of operator Bellman inequality). *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two positive invertible operators such that $0 < mI \le A, B \le MI < I$, and Φ be a normalized positive linear map on $\mathbb{B}(\mathcal{H})$. Then for a given $\alpha > 0$,*

$$\alpha\Phi^{1/p}(I - A\nabla_v B) + \beta I \le \Phi((I - A)^{1/p}\nabla_v(I - B)^{1/p}),$$

where $p > 1$, $v \in [0, 1]$ and

$$\beta = \min_{t\in[m,M]} \left\{ \frac{(1 - M)^{1/p} - (1 - m)^{1/p}}{M - m} t + \frac{M(1 - m)^{1/p} - m(1 - M)^{1/p}}{M - m} - \alpha(1 - t)^{1/p} \right\}.$$

By choosing appropriate α and β in Corollary 6.3.1, we find the following simpler forms.

Corollary 6.3.2. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two positive invertible operators such that $0 < mI \le A, B \le MI < I$, and Φ be a normalized positive linear map on $\mathbb{B}(\mathcal{H})$. Then*

$$\alpha\Phi^{1/p}(I - A\nabla_v B) \le \Phi\big((I - A)^{1/p}\nabla_v (I - B)^{1/p}\big),$$

where $p > 1$, $v \in [0, 1]$ and

$$\alpha = \min_{t\in[m,M]}\left\{\frac{1}{(1-t)^{1/p}}\left(\frac{(1-M)^{1/p} - (1-m)^{1/p}}{M-m}t + \frac{M(1-m)^{1/p} - m(1-M)^{1/p}}{M-m}\right)\right\}.$$

In addition, we have

$$\Phi^{1/p}(I - A\nabla_v B) + \beta I \le \Phi\big((I - A)^{1/p}\nabla_v (I - B)^{1/p}\big),$$

where

$$\beta = \min_{t\in[m,M]}\left\{\frac{(1-M)^{1/p} - (1-m)^{1/p}}{M-m}t + \frac{M(1-m)^{1/p} - m(1-M)^{1/p}}{M-m} - (1-t)^{1/p}\right\}.$$

As an application of Corollary 6.3.2, we have the following scalar Bellman-type inequality.

Corollary 6.3.3. *For $i = 1, 2, \ldots, n$, let a_i, b_i be positive numbers satisfying $0 < m \le a_i$, $b_i \le M < 1$ for some scalars m and M. Then we have for $p > 1$ and $q \le 1$,*

$$\alpha 2^{1-q/p}\sum_{i=1}^{n}\big(2^q - (a_i + b_i)^q\big)^{1/p} \le \sum_{i=1}^{n}\big\{\big(1 - a_i^q\big)^{1/p} + \big(1 - b_i^q\big)^{1/p}\big\}, \tag{6.3.7}$$

where α is as in Corollary 6.3.2.

Proof. For the given a_i, b_i, define the $n \times n$ matrices $A = \operatorname{diag}(a_i^q)$ and $B = \operatorname{diag}(b_i^q)$. Applying the first inequality of Corollary 6.3.2 with $v = 1/2$, we get

$$\alpha(I - A\nabla B)^{1/p} \le (I - A)^{1/p}\nabla(I - B)^{1/p},$$

where we have chosen Φ to be the identity map. In particular, it follows that

$$\alpha\big\|(I - A\nabla B)^{1/p}\big\|_u \le \big\|(I - A)^{1/p}\nabla(I - B)^{1/p}\big\|_u$$

for any unitarily invariant norm $\|\cdot\|_u$. Using the trace norm $\|\cdot\|_1$ as a special case, we obtain

$$\alpha\sum_{i=1}^{n} s_i\big((I - A\nabla B)^{1/p}\big) \le \sum_{i=1}^{n} s_i\big((I - A)^{1/p}\nabla(I - B)^{1/p}\big),$$

where s_i is the ith singular value. This implies

$$\alpha\sum_{i=1}^{n}\left(1-a_i^q\nabla b_i^q\right)^{1/p}\le\sum_{i=1}^{n}\left\{\left(1-a_i^q\right)^{1/p}\nabla\left(1-b_i^q\right)^{1/p}\right\}.$$

That is, noting concavity of the mapping $t\mapsto t^q$,

$$\begin{aligned}\frac{1}{2}\sum_{i=1}^{n}\left\{\left(1-a_i^q\right)^{1/p}+\left(1-b_i^q\right)^{1/p}\right\}&\ge\alpha\sum_{i=1}^{n}\left(1-\frac{a_i^q+b_i^q}{2}\right)^{1/p}\ge\alpha\sum_{i=1}^{n}\left(1-\left(\frac{a_i+b_i}{2}\right)^q\right)^{1/p}\\&=\frac{\alpha}{2^{q/p}}\sum_{i=1}^{n}\left(2^q-(a_i+b_i)^q\right)^{1/p},\end{aligned}$$

which completes the proof. □

Corollary 6.3.4. *Let $A,B\in\mathbb{B}(\mathcal{H})$ be two positive operators such that $mI\le A,B\le MI$ with $0<m<M$. If $f:[m,M]\to[0,\infty)$ is a concave function and $v\in[0,1]$, then the ratio-type inequality*

$$f(A\nabla_v B)\le\frac{1}{\alpha}(f(A)\nabla_v f(B))\tag{6.3.8}$$

holds, where $\alpha=\min_{t\in[m,M]}\{\frac{a_f t+b_f}{f(t)}\}$. Additionally, the following difference-type inequality

$$f(A\nabla_v B)+\beta I\le f(A)\nabla_v f(B)\tag{6.3.9}$$

holds, where $\beta=\min_{t\in[m,M]}\{a_f t+b_f-f(t)\}$.

The reverse inequalities in (6.3.8) *and* (6.3.9) *hold when f is a convex function.*

We conclude this section, by presenting the following simple proof of (6.3.2) and some reversed versions.

Proposition 6.3.1. *Let a_k,b_k be positive numbers such that $\sum_{k=1}^{n}a_k^p\le1$ and $\sum_{k=1}^{n}b_k^p\le1$, for $p\in\mathbb{R}$. Then we have for $v\in[0,1]$,*

$$\left(1-\sum_{k=1}^{n}(a_k^p\nabla_v b_k^p)\right)^{1/p}\ge\left(1-\sum_{k=1}^{n}a_k^p\right)^{1/p}\nabla_v\left(1-\sum_{k=1}^{n}b_k^p\right)^{1/p},\quad(p>1)\tag{6.3.10}$$

and

$$\left(1-\sum_{k=1}^{n}(a_k^p\nabla_v b_k^p)\right)^{1/p}\le\left(1-\sum_{k=1}^{n}a_k^p\right)^{1/p}\nabla_v\left(1-\sum_{k=1}^{n}b_k^p\right)^{1/p},\quad(p<1).\tag{6.3.11}$$

Proof. For $v\in[0,1]$, we define the function

$$f(v)=\left(1-\sum_{k=1}^{n}(a_k^p\nabla_v b_k^p)\right)^{\frac{1}{p}}.$$

Since the summands are linear in v, it is readily seen that f is concave if $p > 1$ and is convex if $p < 1$. Then both inequalities follow from concavity/convexity of f. $\square$

Notice that when $p > 1$, the function $x \mapsto x^p$ is convex for $x > 0$. Therefore, $a_k^p \nabla_v b_k^p \geq (a_k \nabla_v b_k)^p$. This observation together with (6.3.10) imply

$$\left(1 - \sum_{k=1}^{n} (a_k \nabla_v b_k)^p\right)^{1/p} \geq \left(1 - \sum_{k=1}^{n} a_k^p\right)^{1/p} \nabla_v \left(1 - \sum_{k=1}^{n} b_k^p\right)^{1/p}, \quad (p > 1).$$

An elaborated proof of this inequality was given in [185] as an application of (6.3.1). Further, in [185], it was shown that this last inequality is equivalent to (6.3.2).

Notice that convexity of the mapping $x \mapsto x^p$, $p > 1$ allowed the passage from $a_k^p \nabla_v b_k^p$ to $(a_k \nabla_v b_k)^p$. Unfortunately, the same logic does not apply for $p < 1$. However, the following is a more elaborated convexity result. The proof follows immediately upon finding the second derivative of the given function.

Proposition 6.3.2. *For the positive numbers $a_k, b_k \leq 1$ and $p \in \mathbb{R}$, define the function*

$$f(v) = \left(1 - \sum_{k=1}^{n} (a_k \nabla_v b_k)^p\right)^{\frac{1}{p}}, \quad 0 \leq v \leq 1.$$

Then f is concave if $p > 1$, while it is convex if $p < 0$.

From this, we have

$$\left(1 - \sum_{k=1}^{n} (a_k \nabla_v b_k)^p\right)^{1/p} \leq \left(1 - \sum_{k=1}^{n} a_k^p\right)^{1/p} \nabla_v \left(1 - \sum_{k=1}^{n} b_k^p\right)^{1/p}, \quad (p < 0).$$

Using this inequality with the similar way to the proof of [185, Theorem 2.5], we obtain the following reverse of (6.3.2).

Corollary 6.3.5 (Reverse of scalar Bellman inequality). *Let a, a_k, b, b_k be positive scalars satisfying $\sum_{k=1}^{n} a_k^p \leq a^p$ and $\sum_{k=1}^{n} b_k^p \leq b^p$, where $p < 0$. Then the following reverse of* (6.3.2) *holds:*

$$\left(a^p - \sum_{k=1}^{n} a_k^p\right)^{1/p} + \left(b^p - \sum_{k=1}^{n} b_k^p\right)^{1/p} \geq \left((a+b)^p - \sum_{k=1}^{n} (a_k + b_k)^p\right)^{1/p}. \tag{6.3.12}$$

Note that the condition on $p < 0$ of (6.3.12) is different from that on $p > 1$ of (6.3.2).

7 Applications to entropy theory

Historically, *entropy* was studied as Clausius entropy in thermodynamics and Boltzmann entropy in statistical physics. In the papers [220], C. E. Shannon introduced information entropy as an unit of information (which is used as a bit nowadays) and then information theory [36] was established, although von Neumann introduced quantum mechanical entropy by a general mathematical form in earlier days [191].

In this chapter, we study the mathematical properties on some entropy and divergence (relative entropy). We give their definitions here. For probability distributions $\mathbf{p} = \{p_1, p_2, \ldots, p_n\}$ and $\mathbf{q} = \{q_1, q_2, \ldots, q_n\}$, **Shannon entropy** $H(\mathbf{p})$ and **relative entropy** $D(\mathbf{p}|\mathbf{q})$ are defined by

$$H(\mathbf{p}) = -\sum_{j=1}^{n} p_j \log p_j, \quad D(\mathbf{p}|\mathbf{q}) = \sum_{j=1}^{n} p_j \log \frac{p_j}{q_j},$$

with convention $0 \log 0 \equiv 0$ and $p_j = 0$ whenever $q_j = 0$ for some j. Shannon entropy is regarded as the weighted arithmetic mean of the **self-information** $-\log p_j$. In [228], C. Tsallis introduced the **Tsallis entropy**, which is one-parameter extension of Shannon entropy, for the analysis of statistical physics. We use the definitions of the Tsallis entropy and the **Tsallis relative entropy**, for example, [86]:

$$H_r(\mathbf{p}) = -\sum_{j=1}^{n} p_j^{1-r} \ln_r p_j, \quad D_r(\mathbf{p}|\mathbf{q}) = \sum_{j=1}^{n} p_j^{1-r} (\ln_r p_j - \ln_r q_j),$$

where $\ln_r$ is called ***r*-logarithmic function** and defined by $\ln_r(x) = \frac{x^r - 1}{r}$ with the parameter $r \in \mathbb{R}$. It is notable that the inverse function of r-logarithmic function is given by ***r*-exponential function** defined as $\exp_r(x) = (1 + rx)^{1/r}$ which comes from $e^x = \lim_{n\to\infty} (1+\frac{x}{n})^n$ putting $r = \frac{1}{n}$. We also have famous quantities, the **Rényi entropy** [205] defined by $R_r(\mathbf{p}) = \frac{1}{r} \log \sum_{j=1}^{n} p_j^{1-r}$ and the **Rényi relative entropy** $R_r(\mathbf{p}|\mathbf{q}) = \frac{1}{r} \log \sum_{j=1}^{n} p_j^{1-r} q_j^r$ as alternative one-parameter extensions. They recover the usual entropy and divergence in the limit $r \to 0$, namely

$$\lim_{r\to 0} H_r(\mathbf{p}) = \lim_{r\to 0} R_r(\mathbf{p}) = H(\mathbf{p})$$

and

$$\lim_{r\to 0} D_r(\mathbf{p}|\mathbf{q}) = \lim_{r\to 0} R_r(\mathbf{p}|\mathbf{q}) = D(\mathbf{p}|\mathbf{q}),$$

since r-logarithmic function $\ln_r$ uniformly converges to the usual logarithmic function log in the limit $r \to 0$.

Analogically, quantum mechanical (noncommutative) entropy and relative entropy and their one-parameter extended quantities were defined by

$$S(\rho) = -\operatorname{Tr}[\rho \log \rho], \quad D(\rho|\sigma) = \operatorname{Tr}[\rho(\log \rho - \log \sigma)] \tag{7.0.1}$$

https://doi.org/10.1515/9783110643473-007

and

$$S_p(\rho) = -\operatorname{Tr}[\rho^{1-p} \ln_p \rho], \quad D_p(\rho|\sigma) = \operatorname{Tr}[\rho^{1-p}(\ln_p \rho - \ln_p \sigma)] \tag{7.0.2}$$

and

$$R_p(\rho) = \frac{1}{p} \log \operatorname{Tr}[\rho^{1-p}], \quad R_p(\rho|\sigma) = \frac{1}{p} \log \operatorname{Tr}[\rho^{1-p}\sigma^p]$$

for **density operators** (which are self-adjoint positive operators with a unit trace) ρ and σ. A sandwiched Rényi relative entropy and its further generalizations are studied in [188] and references therein; we do not treat them (and Rényi entropy, Rényi relative entropy, α-divergence [5] and sandwiched Rényi relative entropy) in this book. See [199] for quantum entropy and [86] for quantum Tsallis relative entropy.

As J. I. Fujii and E. Kamei introduced the relative operator entropy [52] (see also Definition 4.2.2 and it is notable that the similar definition was previously done in [16]) by

$$S(X|Y) = X^{1/2} \log(X^{-1/2} Y X^{-1/2}) X^{1/2}$$

for $X, Y > 0$, the **Tsallis relative operator entropy** $T_p(X|Y)$ and the **Furuta parametric relative operator entropy** $S_p(X|Y)$ for $p \in \mathbb{R}$ with $p \neq 0$ were defined [240, 93] by

$$T_p(X|Y) = X^{1/2} \ln_p(X^{-1/2} Y X^{-1/2}) X^{1/2} \tag{7.0.3}$$

and

$$S_p(X|Y) = X^{1/2}(X^{-1/2} Y X^{-1/2})^p \log(X^{-1/2} Y X^{-1/2}) X^{1/2}. \tag{7.0.4}$$

We see that $\lim_{p\to 0} T_p(X|Y) = \lim_{p\to 0} S_p(X|Y) = S(X|Y)$. We also find the above three relative operator entropies $S(X|Y)$, $T_p(X|Y)$ and $S_p(X|Y)$ are special forms of the perspective defined in (5.0.1) with log, $\ln_p$ and $x^p \log$. In this section, we mainly study to estimate the bounds on the relative operator entropies.

7.1 Some inequalities on the quantum entropies

In this section, we first recall the quantum Tsallis relative entropy defined in (7.0.2) is written by $D_p(\rho|\sigma) = \frac{\operatorname{Tr}[\rho - \rho^{1-p}\sigma^p]}{p}$ without the function $\ln_p$. We also note that $D_p(\cdot|\cdot)$ may be defined for $A, B > 0$ without the condition $\operatorname{Tr}[A] = \operatorname{Tr}[B] = 1$. We often use this formula generally.

By the use of spectral decompositions of ρ and σ and the scalar inequalities $\frac{1-x^{-t}}{t} \leq \log x \leq \frac{x^t - 1}{t}$ for $x > 0$ and $t > 0$, we have the following relations.

Proposition 7.1.1 (Ruskai–Stillinger [207, 199]). *For the density operators ρ and σ, we have*

(i) $D_{1-p}(\rho|\sigma) \le D(\rho|\sigma) \le D_{1+p}(\rho|\sigma)$, $(0 < p \le 1)$.
(ii) $D_{1+p}(\rho|\sigma) \le D(\rho|\sigma) \le D_{1-p}(\rho|\sigma)$, $(-1 \le p < 0)$.

As an application of Golden–Thompson inequality (which is treated in Section 6.1), we have the following inequality which states the relation between quantum Tsallis relative entropy and the Tsallis relative operator entropy defined in (7.0.3).

Theorem 7.1.1 ([86]). *For $0 \le p < 1$ and any strictly positive operators with unit trace ρ and σ, we have*

$$D_p(\rho|\sigma) \le -\operatorname{Tr}[T_p(\rho|\sigma)]. \tag{7.1.1}$$

Proof. From [114, Theorem 3.4], we have

$$\operatorname{Tr}[e^A \sharp_\alpha e^B] \le \operatorname{Tr}[e^{(1-\alpha)A+\alpha B}]$$

for any $p \in [0,1]$ and $A, B > 0$. Putting $A = \log\rho$ and $B = \log\sigma$, we have

$$\operatorname{Tr}[\rho\sharp_p\sigma] \le \operatorname{Tr}[e^{\log\rho^{1-p}+\log\sigma^p}].$$

Since the Golden–Thompson inequality $\operatorname{Tr}[e^{A+B}] \le \operatorname{Tr}[e^A e^B]$ holds for any Hermitian operators A and B, we have

$$\operatorname{Tr}[e^{\log\rho^{1-p}+\log\sigma^p}] \le \operatorname{Tr}[e^{\log\rho^{1-p}}e^{\log\sigma^p}] = \operatorname{Tr}[\rho^{1-p}\sigma^p].$$

Therefore,

$$\operatorname{Tr}[\rho^{1/2}(\rho^{-1/2}\sigma\rho^{-1/2})^p\rho^{1/2}] \le \operatorname{Tr}[\rho^{1-p}\sigma^p]$$

which implies the theorem. □

We have another upper bound of the quantum Tsallis relative entropy for $X, Y > 0$ and $p \in (0,1]$,

$$D_p(X|Y) \le \frac{\operatorname{Tr}[(X-Y)_+]}{p}, \tag{7.1.2}$$

where $A_+ \equiv \frac{1}{2}(A+|A|)$ and $|A| \equiv (A^*A)^{1/2}$ for any operator A. The proof is directly given by the use of the inequality for $A, B > 0$ and $s \in [0,1]$,

$$\operatorname{Tr}[A^s B^{1-s}] \ge \frac{1}{2}\operatorname{Tr}[A+B-|A-B|]$$

by K. M. R. Audenaert et al. in [13].

We give the fundamental properties for quantum Tsallis relative entropy without proofs.

Proposition 7.1.2 ([86]). *For $p \in (0,1]$ and any density operators ρ and σ, the quantum Tsallis relative entropy $D_p(\rho|\sigma)$ has the following properties:*

(i) *(Nonnegativity)* $D_p(\rho|\sigma) \geq 0$.
(ii) *(Pseudoadditivity)* $D_p(\rho_1 \otimes \rho_2|\sigma_1 \otimes \sigma_2) = D_p(\rho_1|\sigma_1) + D_p(\rho_2|\sigma_2) + pD_p(\rho_1|\sigma_1)D_p(\rho_2|\sigma_2)$.
(iii) *(Joint convexity)* $D_p(\sum_j \lambda_j \rho_j | \sum_j \lambda_j \sigma_j) \leq \sum_j \lambda_j D_p(\rho_j|\sigma_j)$.
(iv) *The quantum Tsallis relative entropy is invariant under the unitary transformation U:*

$$D_p(U\rho U^*|U\sigma U^*) = D_p(\rho|\sigma).$$

(v) *(Monotonicity) For any trace-preserving completely positive linear map Φ, $D_p(\Phi(\rho)|\Phi(\sigma)) \leq D_p(\rho|\sigma)$.*

By the **Hölder inequality**,

$$\left|\mathrm{Tr}[XY]\right| \leq \mathrm{Tr}\left[|X|^s\right]^{1/s} \mathrm{Tr}\left[|Y|^t\right]^{1/t} \tag{7.1.3}$$

for $X, Y > 0$ and $s, t \geq 1$ such that $\frac{1}{s} + \frac{1}{t} = 1$, we obtain the following **generalized Peierls–Bogoliubov inequality** for the quantum Tsallis relative entropy.

Proposition 7.1.3 ([86]). *For $X, Y > 0$ and $p \in (0, 1]$,*

$$D_p(X||Y) \geq \frac{\mathrm{Tr}[X] - (\mathrm{Tr}[X])^{1-p}(\mathrm{Tr}[Y])^p}{p}.$$

Taking the limit $p \to 0$, we recover the original **Peierls–Bogoliubov inequality** [15, 119] from Proposition 7.1.3:

$$D(X|Y) \geq \mathrm{Tr}\big[X(\log \mathrm{Tr}[X] - \log \mathrm{Tr}[Y])\big].$$

We also easily find that the right-hand side in inequality of Proposition 7.1.3 is nonnegative, provided $\mathrm{Tr}[X] \geq \mathrm{Tr}[Y]$.

We review the **maximum quantum entropy principle** in Tsallis statistics. See [61] for the maximum *classical* entropy principle in Tsallis statistics.

In quantum system, the expectation value of a Hermitian matrix (an observable) H in a density matrix (a quantum state) $X \in M_{+,1}(n, \mathbb{C})$ is written as $\mathrm{Tr}[XH]$. Here, we consider the p-expectation value $\mathrm{Tr}[X^{1-p}H]$ as a generalization of the usual expectation value. For $-I \leq X \leq I$ and $p \in (-1, 0) \cup (0, 1)$, we denote the p-exponential function by $\exp_p(X) \equiv (I+pX)^{1/p}$. First, we impose the following constraint on the maximization problem of the quantum Tsallis entropy:

$$C_p = \{X \in M_{+,1}(n, \mathbb{C}) : \mathrm{Tr}[X^{1-p}H] = 0\},$$

for a given $H \in M_h(n, \mathbb{C})$. We denote a usual matrix norm by $\|\cdot\|$, namely for $A \in M(n, \mathbb{C})$ and $x \in \mathbb{C}^n$,

$$\|A\| \equiv \max_{\|x\|=1} \|Ax\|.$$

Then we have the following theorem.

Theorem 7.1.2 ([60]). *Let* $Y = Z_p^{-1} \exp_p(-H/\|H\|)$, *where* $Z_p \equiv \mathrm{Tr}[\exp_p(-H/\|H\|)]$ *is often called partition function, for* $H \in M_h(n, \mathbb{C})$ *and* $p \in (-1,0) \cup (0,1)$. *If* $X \in C_p$, *then* $S_p(X) \leq -c_p \ln_p Z_p^{-1}$, *where* $c_p = \mathrm{Tr}[X^{1-p}]$.

Proof. Since $Z_p \geq 0$ and we have $\ln_p(x^{-1}Y) = \ln_p Y + (\ln_p x^{-1})Y^p$ for $Y \geq 0$ and $x > 0$, we calculate

$$\begin{aligned}\mathrm{Tr}[X^{1-p} \ln_p Y] &= \mathrm{Tr}[X^{1-p} \ln_p\{Z_p^{-1} \exp_p(-H/\|H\|)\}] \\ &= \mathrm{Tr}[X^{1-p}\{-H/\|H\| + \ln_p Z_p^{-1}(I - pH/\|H\|)\}] \\ &= \mathrm{Tr}[X^{1-p}\{\ln_p Z_p^{-1} I - Z_p^{-p} H/\|H\|\}] \\ &= c_p \ln_p Z_p^{-1},\end{aligned}$$

since $\ln_p Z_p^{-1} = \frac{Z_p^{-p}-1}{p}$ by the definition of p-logarithmic function $\ln_p(\cdot)$. By the nonnegativity of the quantum Tsallis relative entropy:

$$\mathrm{Tr}[X^{1-p} \ln_p Y] \leq \mathrm{Tr}[X^{1-p} \ln_p X], \tag{7.1.4}$$

we have

$$S_p(X) = -\mathrm{Tr}[X^{1-p} \ln_p X] \leq -\mathrm{Tr}[X^{1-p} \ln_p Y] = -c_p \ln_p Z_p^{-1}. \qquad \square$$

Regarding the result for a slightly changed constraint, see [60]. Here, we give an extension of (4.2.6) without proof.

Theorem 7.1.3 ([60]). *For* $X, Y \geq 0$, $q \geq 1$ *and* $0 < p \leq 1$, *we have*

$$D_p(X|Y) \leq -\mathrm{Tr}[X \ln_p (X^{-q/2} Y^q X^{-q/2})^{1/q}]. \tag{7.1.5}$$

Taking limit $p \to 0$ and putting $Y := Y^{-1}$ (we then assume $Y > 0$) and then putting $q = p$, (7.1.5) recovers (4.2.2).

We obtained the lower bound (7.1.3) of quantum Tsallis relative entropy as the generalized Peierls–Bogoliubov inequality. We give another lower bound of quantum Tsallis relative entropy under a certain assumption.

Theorem 7.1.4 ([63]). *For* $X, Y > 0$ *and* $p \in (0,1]$, *if* $I \leq Y \leq X$, *then we have*

$$D_p(X|Y) \geq \mathrm{Tr}[X^{1-p} \ln_p (Y^{-1/2} X Y^{-1/2})]. \tag{7.1.6}$$

It is notable that the condition $X \geq Y$ assures the nonnegativity of the right-hand side in the inequality (7.1.6). To prove Theorem 7.1.4, we need the following lemmas. Lemma 7.1.1 below is obtained by Corollary 4.2.6.

Lemma 7.1.1 ([75]). *For positive operators* X, Y *and* $\nu \in (0,1]$, *we have*

$$\mathrm{Tr}[\exp_p(X + Y)] \leq \mathrm{Tr}[\exp_p(X) \exp_p(Y)]. \tag{7.1.7}$$

The inequality (7.1.7) can be regarded as a kind of one-parameter **extension of Golden–Thompson inequality** for $X, Y > 0$. Here, we give a slightly different version of a variational expression for the Tsallis relative entropy. It can be proven by the similar way to in [58, Theorem 2.1].

Lemma 7.1.2 ([63]). *For any $p \in (0,1]$ and any $d \in [0,\infty)$, we have the following relations:*
(i) *If $A, Y > 0$, then we have*

$$d \ln_p\left(\frac{\mathrm{Tr}[\exp_p(A + \ln_p Y)]}{d}\right) = \max\{\mathrm{Tr}[X^{1-p}A] - D_p(X|Y) : X \geq 0, \mathrm{Tr}[X] = d\}.$$

(ii) *If $X, B > 0$ with $\mathrm{Tr}[X] = d$, then*

$$D_p(X|\exp_p(B)) = \max\left\{\mathrm{Tr}[X^{1-p}A] - d\ln_p\left(\frac{\mathrm{Tr}[\exp_p(A + B)]}{d}\right) : A \geq 0\right\}.$$

Lemma 7.1.2 can be also regarded as a parametric extension of Lemma 4.2.1 for $A, B > 0$. See [63] for proofs of Lemma 7.1.1, 7.1.2 and Theorem 7.1.4.

Closing this section, we give a remark on quantum Pinsker inequality and its related result. The quantum Pinsker inequality on quantum relative entropy (divergence) and similar one are known (see, e. g., [202] and [32], resp.)

$$D(\rho|\sigma) = \mathrm{Tr}[\rho(\log\rho - \log\sigma)] \geq \frac{1}{2}\,\mathrm{Tr}\,[|\rho - \sigma|]^2$$

and

$$D(\rho|\sigma) \geq -2\log\mathrm{Tr}[\rho^{1/2}\sigma^{1/2}] \geq \mathrm{Tr}\,[\rho^{1/2} - \sigma^{1/2}]^2.$$

They give the lower bounds for the quantum relative entropy. We consider here the **quantum Jeffrey divergence** defined by

$$J(\rho|\sigma) \equiv \frac{1}{2}\{D(\rho|\sigma) + D(\sigma|\rho)\}.$$

It is seen that the quantum f-divergence corresponds to the perspective. Since $J(\rho|\sigma) = J(\sigma|\rho)$, the quantum Jeffrey divergence has a symmetric property. We show the lower bound of the quantum Jeffrey divergence. To do so, we use the following well-known fact. See [202], for example.

Lemma 7.1.3. *Let $\mathcal{E} : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ be a state transformation. For an operator monotone decreasing function $f : [0,\infty) \to \mathbb{R}$, the monotonicity holds:*

$$D_f(\rho|\sigma) \geq D_f(\mathcal{E}(\rho)|\mathcal{E}(\sigma)),$$

where $D_f(\rho|\sigma) \equiv \mathrm{Tr}[\rho f(\Delta)(I)]$ is the quantum f-divergence, with $\Delta_{\sigma,\rho} \equiv \Delta = LR$ is the **relative modular operator** *such as $L(A) = \sigma A$ and $R(A) = A\rho^{-1}$.*

Theorem 7.1.5 ([89]). *The quantum Jeffrey divergence has the following lower bound:*

$$J(\rho|\sigma) \geq d(\rho,\sigma)\log\left(\frac{1+d(\rho,\sigma)}{1-d(\rho,\sigma)}\right).$$

Proof. We first we show that

$$d(\rho,\sigma) := \frac{1}{2}Tr[|\rho-\sigma|] = \|\rho-\sigma\|_1 = \|P-Q\|_1 = \frac{1}{2}\sum_x |P(x)-Q(x)| =: d_{TV}(P,Q).$$

Let $\mathcal{A} = C^*(\rho_1-\rho_2)$ be commutative C^*-algebra generated by $\rho_1-\rho_2$, and set the map $\mathcal{E} : M(n,\mathbb{C}) \to \mathcal{A}$ as trace preserving, conditional expectation. If we take $p_1 = \mathcal{E}(\rho_1)$ and $p_2 = \mathcal{E}(\rho_2)$, then two elements $(\rho_1-\rho_2)_+$ and $(\rho_1-\rho_2)_-$ of Jordan decomposition of $\rho_1-\rho_2$, are commutative functional calculus of $\rho_1-\rho_2$, and we have $p_1-p_2 = \mathcal{E}(\rho_1-\rho_2) = \mathcal{E}((\rho_1-\rho_2)_+ - (\rho_1-\rho_2)_-) = \mathcal{E}((\rho_1-\rho_2)_+) - \mathcal{E}((\rho_1-\rho_2)_-) = (\rho_1-\rho_2)_+ - (\rho_1-\rho_2)_- = \rho_1-\rho_2$ which implies $\|\rho-\sigma\|_1 = \|P-Q\|_1$.

By Lemma 7.1.3, [217, Proposition 4] and $\|\rho-\sigma\|_1 = \|P-Q\|_1$, we have

$$J(\rho|\sigma) \geq J(P|Q) \geq d_{TV}(P,Q)\log\left(\frac{1+d_{TV}(P,Q)}{1-d_{TV}(P,Q)}\right) = d(\rho,\sigma)\log\left(\frac{1+d(\rho,\sigma)}{1-d(\rho,\sigma)}\right).$$

Here, we note that $f(t) = \frac{1}{2}(t-1)\log t$ is operator convex (since $t\log t$ is operator convex and $\log t$ is operator concave on $(0,\infty)$) which is equivalent to operator monotone decreasing and we have $D_{\frac{1}{2}(t-1)\log t}(\rho|\sigma) = J(\rho|\sigma)$, since $(\Delta_{\sigma,\rho}\log\Delta_{\sigma,\rho})(Y) = \sigma\log\sigma(Y)\rho^{-1} - \sigma\rho^{-1}\log\rho(Y)$. □

7.2 Karamata inequalities and applications

We start by recalling some well-known notions which will be used in this section. Let $\mathbf{x} = (x_1,\ldots,x_n)$ and $\mathbf{y} = (y_1,\ldots,y_n)$ be two finite sequences of real numbers, and let $x_1 \geq \cdots \geq x_n$, $y_1 \geq \cdots \geq y_n$ denote the components of $\mathbf{x}$ and $\mathbf{y}$ in decreasing order, respectively. The n-tuple $\mathbf{y}$ is said to majorize $\mathbf{x}$ (or $\mathbf{x}$ is to be majorized by $\mathbf{y}$) in symbols $\mathbf{x} \prec \mathbf{y}$, if

$$\sum_{i=1}^k x_i \leq \sum_{i=1}^k y_i \quad \text{holds for } k = 1,\ldots,n-1, \quad \text{and} \quad \sum_{i=1}^n x_i = \sum_{i=1}^n y_i.$$

We start with an elegant result as part of the motivation for this section.

Theorem 7.2.1 (Karamata inequality [125]). *Let $f : J \to \mathbb{R}$ be a convex function and $\mathbf{x} = (x_1,\ldots,x_n)$, $\mathbf{y} = (y_1,\ldots,y_n)$ be two n-tuples such that $x_i, y_i \in J$ ($i = 1,\ldots,n$). Then*

$$\mathbf{x} \prec \mathbf{y} \Leftrightarrow \sum_{i=1}^n f(x_i) \leq \sum_{i=1}^n f(y_i).$$

The following extension of majorization theorem is due to Fuchs [48].

Theorem 7.2.2 ([48]). *Let $f : J \to \mathbb{R}$ be a convex function, $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$ two decreasing n-tuples such that $x_i, y_i \in J$ $(i = 1, \dots, n)$ and $\mathbf{p} = (p_1, \dots, p_n)$ be a real n-tuple such that*

$$\sum_{i=1}^{k} p_i x_i \le \sum_{i=1}^{k} p_i y_i \quad (k = 1, \dots, n-1) \quad \textit{and} \quad \sum_{i=1}^{n} p_i x_i = \sum_{i=1}^{n} p_i y_i. \tag{7.2.1}$$

Then

$$\sum_{i=1}^{n} p_i f(x_i) \le \sum_{i=1}^{n} p_i f(y_i). \tag{7.2.2}$$

The conditions (7.2.1) are often called ***p*-majorization** [33]. See also [151, Section A in Chapter 14]. As an application of Theorem 7.2.2, we have the following inequality related to mth moment:

$$\sum_{i=1}^{n} p_i (x_i - \overline{x})^m \le \sum_{i=1}^{n} p_i (y_i - \overline{y})^m,$$

where $\overline{x} = \sum_{i=1}^{n} p_i x_i$ and $\overline{y} = \sum_{i=1}^{n} p_i y_i$. Indeed, this case satisfies the conditions (7.2.1) and the function $f(x) = (x - \overline{x})^m$ is convex for $m \ge 1$.

In this section, we show a **complementary inequality to Karamata inequality**. Let f be a convex function on the interval $[m, M]$, and let $(A_1, \dots, A_n)$, $(B_1, \dots, B_n)$ be two n-tuples of self-adjoint operators with $mI \le A_i, B_i \le MI$ $(i = 1, \dots, n)$ and let $p_1, \dots, p_n$ be positive numbers with $\sum_{i=1}^{n} p_i = 1$. We prove, among other inequalities, if $\sum_{i=1}^{n} p_i A_i = \sum_{i=1}^{n} p_i B_i$, then for a given $\alpha \ge 0$,

$$\sum_{i=1}^{n} p_i f(A_i) \le \beta + \alpha \sum_{i=1}^{n} p_i f(B_i),$$

where $\beta = \max_{t \in [m,M]} \{a_f t + b_f - \alpha f(t)\}$ with $a_f = \frac{f(M)-f(m)}{M-m}$, $b_f = \frac{Mf(m)-mf(M)}{M-m}$. Some applications and remarks are given as well. The following notation is used throughout this book:

$$a_f = \frac{f(M) - f(m)}{M - m} \quad \text{and} \quad b_f = \frac{Mf(m) - mf(M)}{M - m}.$$

We now state our first result in this section.

Theorem 7.2.3 ([81]). *Let $f : [m, M] \to \mathbb{R}$ be a convex function, $x_i, y_i \in [m, M]$ $(i = 1, \dots, n)$, and $p_1, \dots, p_n$ be positive numbers with $\sum_{i=1}^{n} p_i = 1$. If $\sum_{i=1}^{n} p_i x_i = \sum_{i=1}^{n} p_i y_i$, then for a given $\alpha \ge 0$,*

$$\sum_{i=1}^{n} p_i f(y_i) \le \beta + \alpha \sum_{i=1}^{n} p_i f(x_i), \tag{7.2.3}$$

where

$$\beta = \max_{t\in[m,M]}\{a_f t + b_f - \alpha f(t)\}. \tag{7.2.4}$$

Proof. Since f is a convex function, for any $t \in [m, M]$ we can write

$$f(t) \le a_f t + b_f. \tag{7.2.5}$$

By choosing $t = y_i$ in (7.2.5), we get

$$f(y_i) \le a_f y_i + b_f. \tag{7.2.6}$$

Multiplying (7.2.6) by p_i ($i = 1, \dots, n$) and then summing over i from 1 to n, we have

$$\sum_{i=1}^{n} p_i f(y_i) \le a_f \sum_{i=1}^{n} p_i y_i + b_f.$$

Where for a given $\alpha \ge 0$,

$$\begin{aligned}
\sum_{i=1}^{n} p_i f(y_i) - \alpha \sum_{i=1}^{n} p_i f(x_i) &\le a_f \sum_{i=1}^{n} p_i y_i + b_f - \alpha \sum_{i=1}^{n} p_i f(x_i)\\
&\le a_f \sum_{i=1}^{n} p_i y_i + b_f - \alpha f\left(\sum_{i=1}^{n} p_i x_i\right) \qquad (7.2.7)\\
&= a_f \sum_{i=1}^{n} p_i y_i + b_f - \alpha f\left(\sum_{i=1}^{n} p_i y_i\right) \qquad (7.2.8)\\
&\le \max_{t\in[m,M]}\{a_f t + b_f - \alpha f(t)\},
\end{aligned}$$

where in (7.2.7) we used the Jensen inequality, and (7.2.8) follows from the fact that $\sum_{i=1}^{n} p_i x_i = \sum_{i=1}^{n} p_i y_i$. □

Remark 7.2.1.

(i) It is worth emphasizing that we have not used the condition $\sum_{i=1}^{k} p_i x_i \le \sum_{i=1}^{k} p_i y_i$ ($k = 1, \dots, n-1$) in Theorem 7.2.3.

(ii) If one chooses $x_1 = x_2 = \cdots = x_n = \sum_{i=1}^{n} p_i y_i$ in Theorem 7.2.3, then we deduce

$$\sum_{i=1}^{n} p_i f(y_i) \le \beta + \alpha f\left(\sum_{i=1}^{n} p_i y_i\right).$$

Actually, Theorem 7.2.3 can be regarded as an extension of the reverse of Jensen inequality.

By choosing appropriate α and β, we obtain the following result.

Corollary 7.2.1. *Let $f : [m,M] \to (0,\infty)$ be a convex function, $x_i, y_i \in [m,M]$ ($i = 1,\ldots,n$), and $p_1,\ldots,p_n$ be positive numbers with $\sum_{i=1}^n p_i = 1$. If $\sum_{i=1}^n p_i x_i = \sum_{i=1}^n p_i y_i$, then*

$$\sum_{i=1}^n p_i f(y_i) \le K(m,M,f)\sum_{i=1}^n p_i f(x_i), \tag{7.2.9}$$

where $K(m,M,f) = \max\{\frac{a_f t + b_f}{f(t)} : t \in [m,M]\}$ which is just same to (5.2.11). *Additionally,*

$$\sum_{i=1}^n p_i f(y_i) \le C(m,M,f) + \sum_{i=1}^n p_i f(x_i), \tag{7.2.10}$$

where $C(m,M,f) = \max\{a_f t + b_f - f(t) : t \in [m,M]\}$.

Remark 7.2.2. We relax the condition $\sum_{i=1}^n p_i x_i = \sum_{i=1}^n p_i y_i$ in Theorem 7.2.3 and Corollary 7.2.1 to $\sum_{i=1}^n p_i x_i \le \sum_{i=1}^n p_i y_i$. But we impose on the monotone decreasingness to the function f. We keep the other conditions as they are. Then we also have the inequalities (7.2.3), (7.2.9) and (7.2.10).

Let $0 < \epsilon \ll 1$. We here set $m = \epsilon$, $M = 1$, $f(t) = -\log t$, $y_i = q_i$ and $x_i = p_i$ in Corollary 7.2.1. Then $K(\epsilon, 1, -\log) = \max_{\epsilon \le t < 1}(\frac{\log\epsilon}{\epsilon-1}\frac{t-1}{\log t}) = \frac{\log\epsilon}{\epsilon-1} > 0$, since the function $\frac{t-1}{\log t}$ is monotone increasing on $t \in (0,1)$ and $\lim_{t\to 1}\frac{t-1}{\log t} = 1$. We also find $C(\epsilon,1,-\log) = \max_{\epsilon\le t\le 1} g(t)$, where $g(t) = \frac{\log\epsilon}{\epsilon-1}(1-t) + \log t$. By easy computations, we have $g'(t) = \frac{1}{t} - \frac{\log\epsilon}{\epsilon-1}$ and $g''(t) = -t^{-2} < 0$, we thus find $g(t)$ takes a maximum value at $t = \frac{\epsilon-1}{\log\epsilon}$ and it is $C(\epsilon,1,-\log) = g(\frac{\epsilon-1}{\log\epsilon}) = -\log\frac{\log\epsilon}{\epsilon-1} + \frac{\log\epsilon}{\epsilon-1} - 1 = \log S(\epsilon)$, where $S(h)$ is Specht ratio given in [223].

Under the assumption $\sum_{i=1}^n p_i^2 \le \sum_{i=1}^n p_i q_i$ with $\sum_{i=1}^n q_i = 1$, we thus obtain the **reverses of information inequality**:

$$-\left(\frac{\epsilon-1}{\log\epsilon}\right)\sum_{i=1}^n p_i\log q_i \le -\sum_{i=1}^n p_i\log p_i, \quad -\sum_{i=1}^n p_i\log q_i - \log S(\epsilon) \le -\sum_{i=1}^n p_i\log p_i. \tag{7.2.11}$$

Both inequalities give the reverse of **information inequality** (Shannon inequality) [36]:

$$H(\mathbf{p}) = -\sum_{i=1}^n p_i\log p_i \le -\sum_{i=1}^n p_i\log q_i. \tag{7.2.12}$$

It may be noted that $\{p_1,p_2,p_3\} = \{\frac{1}{3},\frac{1}{3},\frac{1}{3}\}$ and $\{q_1,q_2,q_3\} = \{\frac{1}{6},\frac{1}{3},\frac{1}{2}\}$ satisfy the condition $\sum_{i=1}^3 p_i^2 = \sum_{i=1}^3 p_i q_i$, and also $\{p_1,p_2,p_3\} = \{\frac{1}{4},\frac{1}{4},\frac{1}{2}\}$ and $\{q_1,q_2,q_3\} = \{\frac{1}{10},\frac{1}{10},\frac{4}{5}\}$ satisfy the condition $\sum_{i=1}^3 p_i^2 < \sum_{i=1}^3 p_i q_i$, for example.

Similarly, we set $m = \epsilon$, $M = 1$, $f(t) = -\log t$, $y_i = p_i$ and $x_i = q_i$ in Corollary 7.2.1. Then we obtain the following inequalities:

$$-\sum_{i=1}^{n} p_i \log p_i \le -\left(\frac{\log\epsilon}{\epsilon-1}\right)\sum_{i=1}^{n} p_i \log q_i, \quad -\sum_{i=1}^{n} p_i \log p_i \le \log S(\epsilon) - \sum_{i=1}^{n} p_i \log q_i, \tag{7.2.13}$$

under the assumption $\sum\limits_{i=1}^{n} p_i q_i \le \sum\limits_{i=1}^{n} p_i^2$. Since $\frac{\log\epsilon}{\epsilon-1} > 1$ for $0 < \epsilon < 1$ and $\log S(\epsilon)$ is decreasing on $0 < \epsilon < 1$ and $\log S(\epsilon) > 0$ with $\lim\limits_{\epsilon\to 0}\log S(\epsilon) = \infty$ and $\lim\limits_{\epsilon\to 1}\log S(\epsilon) = 0$, both inequalities above do not refine the information inequality given in (7.2.12).

In the following, we aim to extend Theorem 7.2.3 to the Hilbert space operators. To do this end, we need the following lemma.

Lemma 7.2.1. *Let $f : J \to \mathbb{R}$ be a convex function, $A_i \in \mathbb{B}(\mathcal{H})_h$, $(i = 1,\dots,n)$ with the spectra in J, and let $\Phi_i : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$, $(i = 1,\dots,n)$ be positive linear maps such that $\sum\limits_{i=1}^{n} \Phi_i(\mathbf{1}_{\mathcal{H}}) = I$. Then for each unit vector $x \in \mathcal{K}$,*

$$f\left(\left\langle \sum_{i=1}^{n} \Phi_i(A_i)x, x\right\rangle\right) \le \left\langle \sum_{i=1}^{n} \Phi_i(f(A_i))x, x\right\rangle. \tag{7.2.14}$$

Proof. It is well known that if f is a convex function on an interval J, then for each point $(s, f(s))$, there exists a real number C_s such that

$$f(s) + C_s(t-s) \le f(t), \quad \text{for all } t \in J. \tag{7.2.15}$$

(Of course, if f is differentiable at s, then $C_s = f'(s)$.) Fix $s \in J$. Since J contains the spectra of the A_i for $i = 1,\dots,n$, we may replace t in the above inequality by A_i, via a functional calculus to get $f(s)I + C_s A_i - C_s sI \le f(A_i)$. Applying the positive linear maps Φ_i and summing on i from 1 to n, this implies

$$f(s)I + C_s\sum_{i=1}^{n}\Phi_i(A_i) - C_s sI \le \sum_{i=1}^{n}\Phi_i(f(A_i)). \tag{7.2.16}$$

The inequality (7.2.16) easily implies, for each unit vector $x \in \mathcal{K}$,

$$f(s) + C_s\left\langle \sum_{i=1}^{n}\Phi_i(A_i)x, x\right\rangle - C_s s \le \left\langle \sum_{i=1}^{n}\Phi_i(f(A_i))x, x\right\rangle. \tag{7.2.17}$$

On the other hand, since $\sum\limits_{i=1}^{n}\Phi_i(I) = I$ we have $\langle\sum\limits_{i=1}^{n}\Phi_i(A_i)x, x\rangle \in J$ with each unit vector $x \in \mathcal{K}$. Therefore, we may replace s by $\langle\sum\limits_{i=1}^{n}\Phi_i(A_i)x, x\rangle$ in (7.2.17). This yields (7.2.14). □

Our promised generalization of Theorem 7.2.3 can be stated as follows.

Theorem 7.2.4. *Let $f : [m, M] \to \mathbb{R}$ be a convex function, $A_i, B_i \in \mathbb{B}(\mathcal{H})_h$, $(i = 1, \ldots, n)$ with the spectra in $[m, M]$, and let $\Phi_i : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$, $(i = 1, \ldots, n)$ be positive linear maps such that $\sum_{i=1}^{n} \Phi_i(I) = I$. If $\sum_{i=1}^{n} \Phi_i(A_i) = \sum_{i=1}^{n} \Phi_i(B_i)$, then for a given $\alpha \geq 0$,*

$$\sum_{i=1}^{n} \Phi_i(f(A_i)) \leq \beta + \alpha \sum_{i=1}^{n} \Phi_i(f(B_i)),$$

where β is defined as in (7.2.4).

Proof. It follows from the inequality (7.2.5) that $f(A_i) \leq a_f A_i + b_f I$. Applying positive linear maps Φ_i and summing, we obtain $\sum_{i=1}^{n} \Phi_i(f(A_i)) \leq a_f \sum_{i=1}^{n} \Phi_i(A_i) + b_f I$. For each unit vector $x \in \mathcal{K}$, we thus have $\left\langle \sum_{i=1}^{n} \Phi_i(f(A_i))x, x \right\rangle \leq a_f \left\langle \sum_{i=1}^{n} \Phi_i(A_i)x, x \right\rangle + b_f$. Where for a given $\alpha \geq 0$,

$$\begin{aligned}
&\left\langle \sum_{i=1}^{n} \Phi_i(f(A_i))x, x \right\rangle - \alpha \left\langle \sum_{i=1}^{n} \Phi_i(f(B_i))x, x \right\rangle \\
&\quad \leq a_f \left\langle \sum_{i=1}^{n} \Phi_i(A_i)x, x \right\rangle + b_f - \alpha f\left(\left\langle \sum_{i=1}^{n} \Phi_i(B_i)x, x \right\rangle \right) \qquad (7.2.18) \\
&\quad = a_f \left\langle \sum_{i=1}^{n} \Phi_i(A_i)x, x \right\rangle + b_f - \alpha f\left(\left\langle \sum_{i=1}^{n} \Phi_i(A_i)x, x \right\rangle \right) \leq \beta, \qquad (7.2.19)
\end{aligned}$$

where for the inequality (7.2.18) we have used Lemma 7.2.1, and the equality in (7.2.19) follows from the fact that $\sum_{i=1}^{n} \Phi_i(A_i) = \sum_{i=1}^{n} \Phi_i(B_i)$. Consequently,

$$\left\langle \sum_{i=1}^{n} \Phi_i(f(A_i))x, x \right\rangle \leq \beta + \alpha \left\langle \sum_{i=1}^{n} \Phi_i(f(B_i))x, x \right\rangle$$

for each unit vector $x \in \mathcal{K}$. □

Remark 7.2.3. If $f : [m, M] \to \mathbb{R}$ is a concave function, then the reverse inequalities are valid in Theorems 7.2.3 and 7.2.4 with $\beta = \min_{t \in [m,M]} \{a_f t + b_f - \alpha f(t)\}$.

Remark 7.2.4. In Theorem 7.2.4, we put $m = 0$, $M = 1$, $f(t) = t \log t$, $n = 1$ and $\Phi_1 = \frac{1}{\dim \mathcal{H}} \mathrm{Tr}$. $\mathrm{Tr} : \mathbb{B}(\mathcal{H}) \to \mathbb{R}$ is a usual trace. Since $\lim_{t \to +0} t \log t = 0$, we use the usual convention $f(0) = 0$ in standard information theory [36]. Then we have $a_{t \log t} = b_{t \log t} = 0$ and $\beta = \max_{0 < t \leq 1} (-\alpha t \log t) = \frac{\alpha}{e}$ by easy computations. Therefore, for two positive operators ρ, σ satisfying $\mathrm{Tr}[\rho] = \mathrm{Tr}[\sigma] = 1$ (then the condition $\sum_{i=1}^{n} \Phi_i(A_i) = \sum_{i=1}^{n} \Phi_i(B_i)$ is trivially satisfied), we have the following interesting inequality under the assumption:

$$\alpha S(\sigma) \leq S(\rho) + \frac{\alpha}{e} \dim \mathcal{H}, \quad (\alpha \geq 0), \tag{7.2.20}$$

where $S(\rho)$ defined in (7.0.1) is **von Neumann entropy** (**quantum entropy**) [195] for a self-adjoint positive operator ρ with unit trace. The inequality recovers the nonnegativity $S(\rho) \geq 0$, which is a fundamental property of von Neumann entropy, by taking $\alpha = 0$. Also we obtain the inequality:

$$\left|S(\rho) - S(\sigma)\right| \leq \frac{\dim \mathcal{H}}{e}, \tag{7.2.21}$$

by taking $\alpha = 1$ in (7.2.20) and performing a replacement ρ and σ. It may be interesting to compare the inequality given (7.2.21) and the weaker version of **Fannes inequality** [195, equation (11.45)]:

$$\left|S(\rho) - S(\sigma)\right| \leq \mathrm{Tr}\left[|\rho - \sigma|\right] \log \dim \mathcal{H} + \frac{1}{e}. \tag{7.2.22}$$

If the dimension of Hilbert space $\mathcal{H}$ is large,then the upper bound of (7.2.22) is trivially tighter than one of (7.2.21). When $\dim \mathcal{H} = 1$, both upper bounds coincide. For example, for the simple case $\mathrm{Tr}[|\rho - \sigma|] = 1$, the upper bound of (7.2.21) is tighter than one of (7.2.22) when $\dim \mathcal{H} \leq 5$, while the upper bound of (7.2.22) is tighter than one of (7.2.21) when $\dim \mathcal{H} \geq 6$. As a conclusion, the inequality (7.2.22) gives tighter upper bound than ours in almost cases. That is, our inequality (7.2.21) gives a refinement for the weaker version of Fannes's inequality for only special cases.

It is notable that a sharp Fannes-type inequality for the von Neumann entropy was given in [12] by K. M. R. Audenaert. The method given in Remark 7.2.4 is applicable to a generalized function in the following.

Remark 7.2.5. We use same setting in Remark 7.2.4 except for the function $f_r(t) = \frac{t - t^{1-r}}{r}$ for $t > 0$ and $0 < r \leq 1$. Note that $\lim_{r \to 0} f_r(t) = t \log t$ so $f_r(t)$ is a parametric generalization of the function $t \log t$ used in Remark 7.2.4. We easily find $f_r''(t) = (1-r)t^{-r-1} \geq 0$ and $a_{f_r} = b_{f_r} = 0$. Then $\beta = \max_{0<t\leq 1} g_{r,\alpha}(t)$, where $g_{r,\alpha}(t) = \frac{\alpha}{r}(t^{1-r} - t)$. By easy computations, we have $g_{r,\alpha}'(t) = \frac{\alpha}{r}\{(1-r)t^{-r} - 1\}$ and $g_{r,\alpha}''(t) = -\alpha(1-r)t^{-r-1} \leq 0$. Thus $g_{r,\alpha}$ takes maximum at $t = (1-r)^{\frac{1}{r}}$ and $\beta = g_{r,\alpha}((1-r)^{\frac{1}{r}}) = \alpha(1-r)^{\frac{1-r}{r}}$. By Theorem 7.2.4, we thus have

$$\alpha S_r(\sigma) \leq S_r(\rho) + \alpha(1-r)^{\frac{1-r}{r}} \dim \mathcal{H}, \quad (\alpha \geq 0, 0 < r \leq 1), \tag{7.2.23}$$

where $S_r(\rho)$ defined in (7.0.2) is often called **quantum Tsallis entropy**. See [58, 60], for example. The inequality (7.2.23) recovers the inequality (7.2.20) in the limit $r \to 0$, since $\lim_{r\to 0} S_r(\rho) = S(\rho)$ and $\lim_{r \to 0}(1-r)^{\frac{1-r}{r}} = \frac{1}{e}$. Thus we have the nonnegativity $S_r(\rho) \geq 0$ by taking $\alpha = 0$ and the inequality:

$$\left|S_r(\rho) - S_r(\sigma)\right| \leq (1-r)^{\frac{1-r}{r}} \dim \mathcal{H} \tag{7.2.24}$$

by taking $\alpha = 1$ and performing a replacement ρ and σ. The inequality (7.2.24) recovers the inequality (7.2.21), taking the limit of $r \to 0$.

See [88] for a generalized Fannes inequality using a parameter of quantum Tsallis entropy $S_r(\rho)$.

7.3 Relative operator entropy

The relation between relative operator entropy $S(A|B)$ and Tsallis relative operator entropy $T_p(A|B)$ was considered in [87, 246], as follows:

$$A - AB^{-1}A \le T_{-p}(A|B) \le S(A|B) \le T_p(A|B) \le B - A. \tag{7.3.1}$$

The main result of the present section is a set of bounds that are complementary to (7.3.1). Some of our inequalities improve well-known ones. We also prove a reverse inequality involving Tsallis relative operator entropy $T_p(A|B)$.

7.3.1 Refined bounds for relative operator entropies

An important ingredient in our approach is the following. Let f be a convex function on $[a,b] \subseteq \mathbb{R}$. Then the well-known **Hermite–Hadamard inequality** can be expressed as

$$f\left(\frac{a+b}{2}\right) \le \frac{1}{b-a}\int_a^b f(t)\,dt \le \frac{f(a)+f(b)}{2}. \tag{7.3.2}$$

The following theorem gives a considerable refinement of (7.3.1).

Theorem 7.3.1 ([176]). *For any invertible positive operator A and B such that $A \le B$, and $p \in (0,1]$, we have*

$$\begin{aligned} A^{\frac{1}{2}}\left(\frac{A^{-\frac{1}{2}}BA^{-\frac{1}{2}}+I}{2}\right)^{p-1}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}-I)A^{\frac{1}{2}} &\le T_p(A|B) \\ &\le \frac{1}{2}(A\sharp_p B - A\natural_{p-1}B + B - A). \end{aligned} \tag{7.3.3}$$

Proof. Consider the function $f(t) = t^{p-1}$, $p \in (0,1]$. It is easy to check that $f(t)$ is convex on $[1,\infty)$. Bearing in mind the fact

$$\int_1^x t^{p-1}\,dt = \frac{x^p-1}{p},$$

and utilizing the left-hand side of the Hermite–Hadamard inequality, one can see that

$$\left(\frac{x+1}{2}\right)^{p-1}(x-1) \le \frac{x^p-1}{p}, \tag{7.3.4}$$

where $x \ge 1$ and $p \in (0,1]$. On the other hand, it follows from the right-hand side of the Hermite-Hadamard inequality that

$$\frac{x^p-1}{p} \le \left(\frac{x^{p-1}+1}{2}\right)(x-1), \tag{7.3.5}$$

for each $x \geq 1$ and $p \in (0,1]$. Replacing x by $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$ in (7.3.4) and (7.3.5), and multiplying $A^{\frac{1}{2}}$ on both sides, we get the desired result (7.3.3). □

Remark 7.3.1. Simple calculation gives for all $x \geq 1$ and $p \in (0,1]$,

$$0 \leq 1 - \frac{1}{x} \leq \left(\frac{x+1}{2}\right)^{p-1}(x-1) \leq \frac{x^p-1}{p} \leq \left(\frac{x^{p-1}+1}{2}\right)(x-1) \leq x-1, \tag{7.3.6}$$

which means

$$\begin{aligned}0 \leq A - AB^{-1}A &\leq A^{\frac{1}{2}}\left(\frac{A^{-\frac{1}{2}}BA^{-\frac{1}{2}}+I}{2}\right)^{p-1}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}-I)A^{\frac{1}{2}}\\ &\leq T_p(A|B) \leq \frac{1}{2}(A\sharp_p B - A\natural_{p-1}B + B - A) \leq B - A,\end{aligned}$$

for $0 < A \leq B$ and $p \in (0,1]$. Therefore, our inequalities (7.3.3) improve the inequalities (7.3.1) for the case $0 < A \leq B$.

Concerning (7.3.6), the readers may have interests in the following estimations [68, Lemma 1].

(i) For $p \in [1/2,1]$, $\alpha \in [-1,0) \cup (0,1]$ and $x \geq 1$, we have

$$0 \leq 1 - \frac{1}{x} \leq \frac{x^{-p}-1}{-p} \leq \frac{2(x-1)}{x+1} \leq \frac{2(x^\alpha-1)}{\alpha(x^\alpha+1)} \leq \log x \leq \frac{x^\alpha-1}{\alpha x^{\alpha/2}} \leq \frac{x-1}{\sqrt{x}} \leq \frac{x^p-1}{p} \leq x-1.$$

(ii) For $p \in [1/2,1]$, $\alpha \in [-1,0) \cup (0,1]$ and $0 < x \leq 1$, we have

$$1 - \frac{1}{x} \leq \frac{x^{-p}-1}{-p} \leq \frac{x-1}{\sqrt{x}} \leq \frac{x^\alpha-1}{\alpha x^{\alpha/2}} \leq \log x \leq \frac{2(x^\alpha-1)}{\alpha(x^\alpha+1)} \leq \frac{2(x-1)}{x+1} \leq \frac{x^p-1}{p} \leq x-1 \leq 0.$$

See [68, Remark 2] for the case $0 < p < 1/2$.

Proposition 7.3.1. *For $x \geq 1$ and $\frac{1}{2} \leq p \leq 1$,*

$$\frac{x-1}{\sqrt{x}} \leq \left(\frac{x+1}{2}\right)^{p-1}(x-1). \tag{7.3.7}$$

Proof. In order to prove (7.3.7), we set the function $f_p(x) = (\frac{x+1}{2})^{p-1} - \frac{1}{\sqrt{x}}$ where $x \geq 1$ and $\frac{1}{2} \leq p \leq 1$. Since $\frac{df_p(x)}{dp} = (\frac{x+1}{2})^{p-1}\ln(\frac{x+1}{2})$, $\frac{df_p(x)}{dp} \geq 0$ for $x \geq 1$. Thus, we have $f_p(x) \geq f_{1/2}(x) = \frac{\sqrt{2x}-\sqrt{x+1}}{\sqrt{x(x+1)}} \geq 0$ for $x \geq 1$. Therefore, we have the inequality (7.3.7). □

Remark 7.3.2. The first inequality (7.3.3) gives tight lower bound for the Tsallis relative operator entropy $T_p(A|B)$ more than the eighth inequality in [68, Theorem 2.8 (i)], due to Proposition 7.3.1.

Corollary 7.3.1. *For* $0 < B \le A$ *and* $p \in (0,1]$, *we have*

$$\begin{aligned} A\sharp_p B - A\natural_{p-1}B &\le \frac{1}{2}(A\sharp_p B - A\natural_{p-1}B + B - A) \le T_p(A|B) \\ &\le A^{\frac{1}{2}}\left(\frac{A^{-\frac{1}{2}}BA^{-\frac{1}{2}} + I}{2}\right)^{p-1}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}} - I)A^{\frac{1}{2}} \le A\natural_{p+1}B - A\sharp_p B \le 0. \end{aligned}$$

Proof. Put $t = \frac{1}{x} \le 1$ in the inequalities (7.3.6). □

Remark 7.3.3. By numerical computations, we do not have ordering between $(\frac{x^{p-1}+1}{2}) \times (x-1)$ and $\frac{2(x-1)}{x+1}$ for $0 < x \le 1$ and $\frac{1}{2} \le p \le 1$ so that there is no ordering between the second inequality in Corollary 7.3.1 and the sixth inequality in [68, Theorem 2.8(ii)].

Theorem 7.3.2 ([176]). *For* $0 < A \le B$ *and* $p \in (0,1]$, *we have the following inequalities:*

$$L_p(A,B) + K_p(A,B) \le T_p(A|B) \le R_p(A,B) + K_p(A,B), \tag{7.3.8}$$

and

$$J_p(A,B) - 2R_p(A,B) \le T_p(A|B) \le J_p(A,B) - 2L_p(A,B), \tag{7.3.9}$$

where

$$\begin{aligned} K_p(A,B) &= A^{1/2}\left(\frac{A^{-1/2}BA^{-1/2} + I}{2}\right)^{p-1}(A^{-1/2}BA^{-1/2} - I)A^{1/2}, \\ J_p(A,B) &= \frac{1}{2}(A\sharp_p B - A\natural_{p-1}B + B - A), \\ L_p(A,B) &= \frac{1}{24}(p-1)(p-2)(A\sharp_p B - 3A\natural_{p-1}B + 3A\natural_{p-2}B - A\natural_{p-3}B), \\ R_p(A,B) &= \frac{1}{24}(p-1)(p-2)(A\natural_3 B - 3A\natural_2 B + 3B - A). \end{aligned}$$

Proof. According to [193, Theorem 1](or [194, Lemma 1.10.4]), if $f : [a,b] \to \mathbb{R}$ is a twice differentiable function that there exists real constants m and M so that $m \le f'' \le M$, then

$$m\frac{(b-a)^2}{24} \le \frac{1}{b-a}\int_a^b f(t)\,dt - f\left(\frac{a+b}{2}\right) \le M\frac{(b-a)^2}{24}, \tag{7.3.10}$$

$$m\frac{(b-a)^2}{12} \le \frac{f(a)+f(b)}{2} - \frac{1}{b-a}\int_a^b f(t)\,dt \le M\frac{(b-a)^2}{12}. \tag{7.3.11}$$

Putting $f(t) = t^{p-1}$ with $p \in (0,1]$ and $a = 1$, $b = x$ in the above inequalities, then we have the desired results by a similar way to the proof of Theorem 7.3.1. □

Remark 7.3.4. The first inequality of (7.3.8) and the second inequality of (7.3.9) give tighter bounds of Tsallis relative entropy $T_p(A|B)$ than those in the inequalities (7.3.3), because of the following reasons:

(i) The first inequality of (7.3.10) gives tight lower bound more than the first inequality of the Hermite–Hadamard's inequality (7.3.2).
(ii) The first inequality of (7.3.11) gives tight upper bound more than the second inequality of the Hermite–Hadamard's inequality (7.3.2).

As we mentioned in (7.3.1), we have the inequalities [87] as a part of (7.3.1):

$$A - AB^{-1}A \le T_p(A|B) \le B - A, \tag{7.3.12}$$

for $A, B > 0$ and $p \in [-1,0) \cup (0,1]$ and $\lim_{p\to 0} T_p(A|B) = S(A|B)$. In the sequel, we give precise bounds related to (7.3.12). We start with the following known properties of the Tsallis relative operator entropy. See [120, Theorem 1] or [121, Theorem 2.5(ii)] for example.

Proposition 7.3.2. *For $A, B > 0$ and $p, q \in [-1,0) \cup (0,1]$ with $p \le q$, we have*

$$T_p(A|B) \le T_q(A|B).$$

This proposition can be proven by the monotone increasingness of $\frac{x^p-1}{p}$ on $p \in [-1,0) \cup (0,1]$ for any $x > 0$, and implies the following inequalities (which include the inequalities (7.3.12)) [240]:

$$A - AB^{-1}A = T_{-1}(A|B) \le T_{-p}(A|B) \le S(A|B) \le T_p(A|B) \le T_1(A|B) = B - A$$

for $A, B > 0$ and $p \in (0,1]$. The general results were established in [197] by the notion of perspective functions. In addition, the interesting and significant results for relative operator entropy were given in [44] for the case $B \ge A$. In this subsection, we treat the relations on the Tsallis relative operator entropy under the assumption such that $0 < A \le B$ or $0 < B \le A$.

The inequalities given in Theorem 7.3.1 of Subsection 7.3.1 are improvements of the inequalities (7.3.12). In the present subsection, we give the alternative bounds for the Tsallis relative operator entropy. The condition $A \le B$ in Theorem 7.3.1 can be modified by $uA \le B \le vA$ with $u \ge 1$ so that we use this style (which is often called a sandwich condition) in the present section. Note that the condition $uA \le B \le vA$ with $u \ge 1$ includes the condition $A \le B$ as a special case, also the condition $uA \le B \le vA$ with $v \le 1$ includes the condition $B \le A$ as a special case.

Theorem 7.3.3 ([78]). *For $A, B > 0$ such that $uA \le B \le vA$ with $u, v > 0$ and $-1 \le p \le 1$ with $p \ne 0$, we have*

$$S_{p/2}(A|B) \le T_p(A|B) \le \frac{S(A|B) + S_p(A|B)}{2}, \quad (u \ge 1). \tag{7.3.13}$$

If $v \le 1$, then the reverse inequalities in (7.3.13) hold.

Proof. For $x \geq 1$ and $-1 \leq p \leq 1$ with $p \neq 0$, we define the function $f(t) = x^{pt} \log x$ on $0 \leq t \leq 1$. Since $\frac{d^2 f(t)}{dt^2} = p^2 x^{pt} (\log x)^3 \geq 0$ for $x \geq 1$, the function $f(t)$ is convex on t, for the case $x \geq 1$. Thus we have

$$x^{p/2} \log x \leq \frac{x^p - 1}{p} \leq \left(\frac{x^p + 1}{2} \right) \log x \tag{7.3.14}$$

by Hermite–Hadamard inequality, since $\int_0^1 f(t)\,dt = \frac{x^p - 1}{p}$. Note that $I \leq uI \leq A^{-1/2} B A^{-1/2} \leq vI$ from the condition $u \geq 1$. By **Kubo–Ando theory** [137], it is known that for the **representing function** $f_m(x) = 1mx$ for operator mean m, the scalar inequality $f_m(x) \leq f_n(x)$, $(x > 0)$ is equivalent to the operator inequality $AmB \leq AnB$ for all strictly positive operators A and B. Thus we have the inequality

$$\begin{aligned}
&A^{1/2}(A^{-1/2}BA^{-1/2})^{p/2} \log(A^{-1/2}BA^{-1/2})A^{1/2} \\
&\leq \frac{A \sharp_p B - A}{p} \\
&\leq \frac{A^{1/2} \log(A^{-1/2}BA^{-1/2})A^{1/2} + A^{1/2}(A^{-1/2}BA^{-1/2})^p \log(A^{-1/2}BA^{-1/2})A^{1/2}}{2}
\end{aligned}$$

by the inequality (7.3.14). The reverse inequalities for the case $v \leq 1$ can be similarly shown by the concavity of the function $f(t)$ on t, for the case $0 < x \leq 1$, taking into account the condition $0 < uI \leq A^{-1/2}BA^{-1/2} \leq vI \leq I$. □

We note that both side in the inequalities (7.3.13) and their reverses converge to $S(A|B)$ in the limit $p \to 0$. From the proof of Theorem 7.3.3, for strictly positive operators A and B, we see the following interesting relation between the Tsallis relative operator entropy $T_p(A|B)$ and the generalized relative operator entropy $S_p(A|B)$,

$$\int_0^1 S_{pt}(A|B)\,dt = T_p(A|B).$$

Remark 7.3.5. Let A and B be strictly positive operators such that $uA \leq B \leq vA$ with $u, v > 0$ and let $-1 \leq p \leq 1$ with $p \neq 0$. For the case $0 < p \leq 1$ and $u \geq 1$, we see

$$S(A|B) \leq S_{p/2}(A|B) \leq T_p(A|B) \leq \frac{S(A|B) + S_p(A|B)}{2} \leq S_p(A|B)$$

from the inequalities (7.3.13), since $x^p \log x$ is monotone increasing on $0 < p \leq 1$ and $(\frac{x^p+1}{2}) \log x \leq x^p \log x$ for $x \geq 1$ and $0 < p \leq 1$. For the case $-1 \leq p < 0$ and $v \leq 1$, we also see that the reverse inequalities hold since $x^p \log x$ is monotone increasing on $-1 \leq p < 0$ and $(\frac{x^p+1}{2}) \log x \geq x^p \log x$ for $0 < x \leq 1$ and $-1 \leq p < 0$.

Remark 7.3.6. We compare the bounds of $\frac{x^p-1}{p}$ in the inequalities (7.3.13) with the result given in [176]:

$$\left(\frac{x+1}{2} \right)^{p-1} (x-1) \leq \frac{x^p - 1}{p} \leq \left(\frac{x^{p-1} + 1}{2} \right)(x-1), \quad (x \geq 1, 0 < p \leq 1).$$

(i) We have no ordering between $x^{p/2}\log x$ and $(\frac{x+1}{2})^{p-1}(x-1)$.
(ii) We have no ordering between $(\frac{x^p+1}{2})\log x$ and $(\frac{x^{p-1}+1}{2})(x-1)$.

Therefore, we claim Theorem 7.3.3 is not trivial result.

We can give further relations on Tsallis relative operator entropy. However, we omit the proof, since it is not so difficult (but complicated) computations. See [78] for the proof.

Theorem 7.3.4 ([78]). *For $A,B>0$ such that $uA\le B\le vA$ with $u\ge 1$ and $-1\le p\le 1$ with $p\ne 0$, we have*

$$\frac{T_p(A|B)-T_{p-1}(A|B)}{2}\le 4\left\{T_p\left(A\Big|\frac{A+B}{2}\right)-T_{p-1}\left(A\Big|\frac{A+B}{2}\right)\right\}$$
$$\le\frac{T_p(A|B)-T_1(A|B)}{p-1}\le\frac{T_p(A|B)-T_{p-1}(A|B)}{2}+\frac{A\natural_2(B-A)}{4}.$$

Remark 7.3.7. By [78, Remark 2.7], we conclude that there is no ordering between Theorem 7.3.4 and Theorem 7.3.1. We thus claim Theorem 7.3.4 is also not a trivial result.

Taking the limit $p\to 0$ in Theorem 7.3.4, we have the following corollary.

Corollary 7.3.2. *For strictly positive operators A and B such that $uA\le B\le vA$ with $u\ge 1$, we have*

$$\frac{S(A|B)-T_{-1}(A|B)}{2}\le 4\left\{S\left(A\Big|\frac{A+B}{2}\right)-T_{-1}\left(A\Big|\frac{A+B}{2}\right)\right\}$$
$$\le T_1(A|B)-S(A|B)\le\frac{S(A|B)-T_{-1}(A|B)}{2}+\frac{A\natural_2(B-A)}{4}.$$

Based on the inequalities obtained in this subsection, in relation to Proposition 7.3.2, we can discuss on monotonicity for the parameter of relative operator entropies and for the weighted parameter of weighted operator means. See [78] for the interested readers.

7.3.2 Reverses to information monotonicity inequality

As we stated in Chapter 2, we have the following **Zuo–Liao inequality** with Kantorovich constant $K(\cdot)$.

Proposition 7.3.3. *For $A,B>$ such that $A<h'A\le B\le hA$ or $0<hA\le B\le h'A<A$, we have*

$$K^r(h')A\sharp_p B\le A\nabla_p B\le K^R(h)A\sharp_p B, \tag{7.3.15}$$

where $p\in[0,1]$, $r=\min\{p,1-p\}$, $R=\max\{p,1-p\}$.

The **Ando inequality** [6, Theorem 3] says that if $A, B > 0$ and Φ is a positive linear map, then

$$\Phi(A\sharp_p B) \le \Phi(A)\sharp_p \Phi(B), \quad p \in [0,1]. \tag{7.3.16}$$

Concerning inequality (7.3.16), we have the following corollary.

Corollary 7.3.3. *For $A, B > 0$ such that $A < h'A \le B \le hA$ or $0 < hA \le B \le h'A < A$, and a positive linear map Φ on $\mathbb{B}(\mathcal{H})$, we have*

$$\begin{aligned}\frac{K^r(h')}{K^R(h)}\Phi(A\sharp_p B) &\le \frac{1}{K^R(h)}\Phi(A\nabla_p B) \le \Phi(A)\sharp_p\Phi(B)\\ &\le \frac{1}{K^r(h')}\Phi(A\nabla_p B) \le \frac{K^R(h)}{K^r(h')}\Phi(A\sharp_p B),\end{aligned} \tag{7.3.17}$$

where $p \in [0,1]$, $r = \min\{p, 1-p\}$ and $R = \max\{p, 1-p\}$.

Proof. If we apply a positive linear map Φ in (7.3.15), we infer

$$K^r(h')\Phi(A\sharp_p B) \le \Phi(A\nabla_p B) \le K^R(h)\Phi(A\sharp_p B). \tag{7.3.18}$$

On the other hand, if we take $A = \Phi(A)$ and $B = \Phi(B)$ in (7.3.15) we can write

$$K^r(h')\Phi(A)\sharp_p\Phi(B) \le \Phi(A\nabla_p B) \le K^R(h)\Phi(A)\sharp_p\Phi(B). \tag{7.3.19}$$

Now, combining inequality (7.3.18) and (7.3.19), we deduce the desired inequalities (7.3.17). □

Remark 7.3.8. By virtue of a generalized Kantorovich constant $K(h, p)$ defined in (2.0.8), in the matrix setting, J.-C. Bourin et al. in [29, Theorem 6] gave the following reverse of Ando inequality for a positive linear map:

For $A, B > 0$ such that $mA \le B \le MA$, and a positive linear map Φ, we have

$$\Phi(A)\sharp_p\Phi(B) \le \frac{1}{K(h,p)}\Phi(A\sharp_p B), \quad p \in [0,1], \tag{7.3.20}$$

where $h = \frac{M}{m}$. The above result naturally extends one proved in Lee [138, Theorem 4] for $p = \frac{1}{2}$.

Of course, the constant $\frac{K^R(h)}{K^r(h')}$ is not better than $\frac{1}{K(\frac{h}{h'},p)}$. Concerning the sharpness of the estimate (7.3.20), see [29, Lemma 7]. However, our bounds on $\Phi(A)\sharp_p\Phi(B)$ are calculated by the original Kantorovich constant $K(h)$ without the generalized one $K(h,p)$. It is also interesting our bounds on $\Phi(A)\sharp_p\Phi(B)$ are expressed by $\Phi(A\nabla_p B)$ with only one constant either h or h'.

After discussion on inequalities related to the operator mean with a positive linear map, we give a result on Tsallis relative operator entropy (which is the main theme in

this section) with a positive linear map. It is well known that Tsallis relative operator entropy enjoys the following **information monotonicity** for $-1 \le p \le 1$ with $p \neq 0$ and a normalized positive linear map Φ on $\mathbb{B}(\mathcal{H})$ (see [87, Proposition 2.3] and [54, Theorem 3.1]):

$$\Phi\big(T_p(A|B)\big) \le T_p\big(\Phi(A)|\Phi(B)\big). \tag{7.3.21}$$

Utilizing (2.1.1), we have the following counterpart of (7.3.21):

Theorem 7.3.5 ([176]). *For $A, B > 0$ and a normalized positive linear map Φ on $\mathbb{B}(\mathcal{H})$, we have*

$$\begin{aligned}&\frac{2r}{p}(\Phi(A\nabla B) - \Phi(A)\sharp\Phi(B)) + T_p(\Phi(A)|\Phi(B))\\&\quad\le \Phi(B-A) \le \frac{2R}{p}(\Phi(A\nabla B) - \Phi(A\sharp B)) + \Phi\big(T_p(A|B)\big),\end{aligned} \tag{7.3.22}$$

where $p \in (0,1]$, $r = \min\{p, 1-p\}$ and $R = \max\{p, 1-p\}$.

If $A, B > 0$, then

$$\operatorname{Tr}[A-B] \le D_p(A|B) \le -\operatorname{Tr}\big[T_p(A|B)\big], \quad p \in (0,1]. \tag{7.3.23}$$

Note that the first inequality of (7.3.23) is due to T. Furuta [96, Proposition F] and the second inequality is due to S. Furuichi et al. [86, Theorem 2.2]. As a direct consequence of Theorem 7.3.5 with $\Phi(X) = \frac{\operatorname{Tr}[X]}{\dim \mathcal{H}}$ for $X \in \mathbb{B}(\mathcal{H})$, we have the following interesting relation by the use of the generalized Peierls–Bogoliubov inequality, Proposition 7.1.3. The first inequality in (7.3.24) below gives a reverse to (7.3.23). The second inequality in (7.3.24) below gives a refinement of the first inequality in (7.3.23).

Corollary 7.3.4. *For $A, B > 0$, we have*

$$\begin{aligned}&\frac{2R}{p}\operatorname{Tr}[A\sharp B - A\nabla B] - \operatorname{Tr}\big[T_p(A|B)\big]\\&\quad\le \operatorname{Tr}[A-B] \le \frac{2r}{p}\big(\operatorname{Tr}[A]\sharp\operatorname{Tr}[B] - \operatorname{Tr}[A]\nabla\operatorname{Tr}[B]\big) + D_p(A|B),\end{aligned} \tag{7.3.24}$$

where $p \in (0,1]$, $r = \min\{p, 1-p\}$ and $R = \max\{p, 1-p\}$.

We give here further reverse inequalities to the information monotonicity inequality given in (7.3.21). T. Furuta proved in [96, Theorem 2.1] the following two reverse inequalities involving Tsallis relative operator entropy $T_p(A|B)$ via the generalized Kantorovich constant $K(h,p)$:

For $A, B > 0$, such that $0 < m_1\mathbf{1}_{\mathcal{H}} \le A \le M_1\mathbf{1}_{\mathcal{H}}$ and $0 < m_2\mathbf{1}_{\mathcal{H}} \le B \le M_2\mathbf{1}_{\mathcal{H}}$ ($\mathbf{1}_{\mathcal{H}}$ represents an identity operator on $\mathcal{H}$), and a normalized positive linear map Φ on $\mathbb{B}(\mathcal{H})$, we have

$$T_p(\Phi(A)|\Phi(B)) \le \left(\frac{1-K(h,p)}{p}\right)\Phi(A)\sharp_p\Phi(B) + \Phi\big(T_p(A|B)\big),$$

and

$$T_p(\Phi(A)|\Phi(B)) \le F(h,p)\Phi(A) + \Phi(T_p(A|B)),$$

where

$$F(h,p) = \frac{m^p}{p}\left(\frac{h^p - h}{h-1}\right)(1 - K(h,p)^{\frac{1}{p-1}}),$$

with $m = \frac{m_2}{M_1}$, $M = \frac{M_2}{m_1}$, $h = \frac{M}{m}$ and $p \in (0,1]$. There are a few other results in this direction; see, for example, [95, 176]. In the sequel, we give alternative bounds for two Furuta's inequalities above.

In Theorem 7.3.22, we presented a counterpart of (7.3.21). We here present alternative reverse inequalities of (7.3.21) and give a kind of reverse to the Löwner–Heinz inequality. Our main idea and technical tool are Lemma 7.3.1 below. Most of the results below are rather straightforward consequences of Lemma 7.3.1. The following lemma plays a crucial role in those proofs.

Lemma 7.3.1. *Let* $0 < m \le 1 \le t \le M$.
(i) *If* $p \le 1$ *with* $p \ne 0$, *then*

$$M^{p-1}(t-1) \le \frac{t^p - 1}{p} \le m^{p-1}(t-1). \tag{7.3.25}$$

(ii) *If* $p \ge 1$, *then*

$$m^{p-1}(t-1) \le \frac{t^p - 1}{p} \le M^{p-1}(t-1). \tag{7.3.26}$$

Proof. Take an interval $J \subset \mathbb{R}$ and assume that $f : J \to \mathbb{R}$ is a continuous differentiable function such that $\alpha \le f'(t) \le \beta$ for any $t \in J$ where $\alpha, \beta \in \mathbb{R}$. It is an evident fact that two functions $g_\alpha(t) = f(t) - \alpha t$ and $g_\beta(t) = \beta t - f(t)$ are monotone increasing functions, that is,

$$a \le b \Rightarrow \begin{cases} g_\alpha(a) \le g_\alpha(b) \\ g_\beta(a) \le g_\beta(b) \end{cases} \Leftrightarrow \begin{cases} f(a) - \alpha a \le f(b) - \alpha b \\ \beta a - f(a) \le \beta b - f(b), \end{cases} \tag{7.3.27}$$

for any $a, b \in [m, M] \subseteq J$ where $0 < m \le M$.

Letting $f(x) = x^p$ with $x \in [m, M]$ and $0 < p \le 1$, a little calculation leads to

$$0 < m \le a \le b \le M \Rightarrow \begin{cases} a^p - pM^{p-1}a \le b^p - pM^{p-1}b \\ pm^{p-1}a - a^p \le pm^{p-1}b - b^p. \end{cases} \tag{7.3.28}$$

Dividing the both sides in two inequalities given in (7.3.28) by a^p and taking $t = \frac{b}{a}$, we get

$$0 < \frac{m}{a} \le 1 \le t \le \frac{M}{a} \Rightarrow \begin{cases} p(\frac{M}{a})^{p-1}(t-1) \le t^p - 1 \\ t^p - 1 \le p(\frac{m}{a})^{p-1}(t-1). \end{cases}$$

Setting $\frac{m}{a}$ and $\frac{M}{a}$ equal to m and M, respectively, we obtain the desired inequalities in (7.3.25). For the case of $p < 0$, since $f(x) = x^p$ is decreasing and we find $\alpha = pm^{p-1}$ and $\beta = pM^{p-1}$ in the setting of $g_\alpha(t) = f(t) - \alpha t$ and $g_\beta(t) = \beta t - f(t)$, we get similarly (7.3.28) which implies (7.3.25).

For the case of $p \geq 1$, since $f(x) = x^p$ is increasing and we find $\alpha = pm^{p-1}$ and $\beta = pM^{p-1}$ in the setting of $g_\alpha(t) = f(t) - \alpha t$ and $g_\beta(t) = \beta t - f(t)$, we get similarly

$$0 < m \leq a \leq b \leq M \Rightarrow \begin{cases} a^p - pm^{p-1}a \leq b^p - pm^{p-1}b \\ pM^{p-1}a - a^p \leq pM^{p-1}b - b^p. \end{cases}$$

This implies the inequalities (7.3.26). □

From the preceding result, one may derive an interesting operator inequality.

Theorem 7.3.6 ([175]). *For $A, B > 0$ such that $0 < mA \leq A \leq B \leq MA$ and a normalized positive linear map Φ on $\mathbb{B}(\mathcal{H})$, we have:*
(i) *If $-1 \leq p \leq 1$ with $p \neq 0$, then*

$$T_p(\Phi(A)|\Phi(B)) \leq \Phi(T_p(A|B)) + (m^{p-1} - M^{p-1})\Phi(B - A). \tag{7.3.29}$$

(ii) *If $1 \leq p \leq 2$, then*

$$\Phi(T_p(A|B)) \leq T_p(\Phi(A)|\Phi(B)) + (M^{p-1} - m^{p-1})\Phi(B - A). \tag{7.3.30}$$

Proof. On account of the first inequality in (7.3.25), we infer that

$$M^{p-1}\Phi(B - A) \leq \Phi(T_p(A|B)),$$

and the second one gives

$$T_p(\Phi(A)|\Phi(B)) \leq m^{p-1}\Phi(B - A).$$

Combining above two inequalities with the information monotonicity inequality (7.3.21) for $-1 \leq p \leq 1$ with $p \neq 0$, we have the desired inequality (7.3.29). In the case of $p \geq 1$, by (7.3.26) we infer that

$$M^{p-1}\Phi(B - A) \geq \Phi(T_p(A|B)), \quad T_p(\Phi(A)|\Phi(B)) \geq m^{p-1}\Phi(B - A).$$

We obtain the inequality (7.3.30), since we have the following relation:

$$\Phi(T_p(A|B)) \geq T_p(\Phi(A)|\Phi(B)), \quad 1 \leq p \leq 2$$

which can be shown in a similar way to the proof of [54, Theorem 2.2]. □

Using the same strategy as in the proof of [176, Corollary 2], we get the following converse of Ando inequality $\Phi(A\sharp B) \leq \Phi(A)\sharp\Phi(B)$, for each parameter.

Corollary 7.3.5. *Let the assumptions of Theorem 7.3.6 be satisfied. Then we have the following inequalities:*

(i) *If* $0 \le p \le 1$,

$$\Phi(A)\sharp_p\Phi(B) \le \Phi(A\sharp_p B) + p(m^{p-1} - M^{p-1})\Phi(B - A).$$

(ii) *If* $-1 \le p \le 0$,

$$\Phi(A)\natural_p\Phi(B) \ge \Phi(A\natural_p B) + p(m^{p-1} - M^{p-1})\Phi(B - A).$$

(iii) *If* $1 \le p \le 2$,

$$\Phi(A\natural_p B) \le \Phi(A)\natural_p\Phi(B) + p(M^{p-1} - m^{p-1})\Phi(B - A).$$

Remark 7.3.9. Let $\Phi(X) = \frac{1}{\dim \mathcal{H}} \mathrm{Tr}[X]$ and let A, B be density operators (which are positive operators with a unit trace). Then Corollary 7.3.5 gives $\mathrm{Tr}[A\sharp_p B] \ge 1$ for $0 \le p \le 1$, and $\mathrm{Tr}[A\natural_p B] \le 1$ for $-1 \le p \le 0$ or $1 \le p \le 2$.

As an immediate consequence of Corollary 7.3.5, we have the following result.

Corollary 7.3.6. *Replacing A by $\mathbf{1}_{\mathcal{H}}$ and B by A in Corollary 7.3.5, we get the following inequalities:*

(i) *If* $0 \le p \le 1$, *then* $\Phi^p(A) \le \Phi(A^p) + p(m^{p-1} - M^{p-1})(\Phi(A) - \mathbf{1}_{\mathcal{H}})$.
(ii) *If* $-1 \le p \le 0$, *then* $\Phi^p(A) \ge \Phi(A^p) + p(m^{p-1} - M^{p-1})(\Phi(A) - \mathbf{1}_{\mathcal{H}})$.
(iii) *If* $1 \le p \le 2$, *then* $\Phi(A^p) \le \Phi^p(A) + p(M^{p-1} - m^{p-1})(\Phi(A) - \mathbf{1}_{\mathcal{H}})$.

We recall the following famous inequality [145, 107].

Theorem 7.3.7 (Löwner–Heinz inequality). *For $A, B > 0$, we have*

$$A \le B \Rightarrow A^p \le B^p, \quad 0 \le p \le 1.$$

It is essential to notice that the Löwner–Heinz inequality does not always hold for $p > 1$. Using Lemma 7.3.1, we get a kind of extension and reverse of the Löwner–Heinz inequality under the assumption $\|A\|\mathbf{1}_{\mathcal{H}} \le B$ (here $\|A\|$ stands for the usual operator norm of A). For the sake of convenience, we cite a useful lemma which we will use in the below.

Lemma 7.3.2. *For $A > 0$, we have:*

(i) *If* $0 \le p \le 1$, *then*

$$A^p \le \|A\|^p\mathbf{1}_{\mathcal{H}} - p\|A\|^{p-1}(\|A\|\mathbf{1}_{\mathcal{H}} - A). \tag{7.3.31}$$

(ii) *If* $p \ge 1$ *or* $p \le 0$, *then*

$$\|A\|^p\mathbf{1}_{\mathcal{H}} - p\|A\|^{p-1}(\|A\|\mathbf{1}_{\mathcal{H}} - A) \le A^p. \tag{7.3.32}$$

Proof. Since $A \le \|A\|\mathbf{1}_{\mathcal{H}}$, $M := \|A\|$ and A commutes with $\mathbf{1}_{\mathcal{H}}$, we can use directly the scalar inequality $a^p - pM^{p-1}a \le b^p - pM^{p-1}b$ for $a \le b$ as $a = A$ and $b = \|A\|\mathbf{1}_{\mathcal{H}}$. Then we can obtain (7.3.31). The inequality (7.3.32) can be proven by the similar way by using the inequality $pm^{p-1}a - a^p \le pm^{p-1}b - b^p$ for $a \le b$. □

We show the following theorem.

Theorem 7.3.8 ([175]). *For $A, B > 0$ such that $\|A\|\mathbf{1}_{\mathcal{H}} \le B$, we have*
(i) *If $0 \le p \le 1$, then*

$$p\|B\|^{p-1}(B - A) \le B^p - A^p. \tag{7.3.33}$$

(ii) *If $p \ge 1$ or $p \le 0$, then*

$$B^p - A^p \le p\|B\|^{p-1}(B - A). \tag{7.3.34}$$

Proof. Replacing a by $\|A\|$ and then applying functional calculus for the operator B in the first inequality in (7.3.28), we get

$$\|A\|\mathbf{1}_{\mathcal{H}} \le B \le M\mathbf{1}_{\mathcal{H}} \Rightarrow (\|A\|^p - pM^{p-1}\|A\|)\mathbf{1}_{\mathcal{H}} \le B^p - pM^{p-1}B.$$

On account of the inequality (7.3.31), we have

$$A^p - pM^{p-1}A \le (\|A\|^p - pM^{p-1}\|A\|)\mathbf{1}_{\mathcal{H}}.$$

This is the same as saying

$$\|A\|\mathbf{1}_{\mathcal{H}} \le B \le M\mathbf{1}_{\mathcal{H}} \Rightarrow A^p - pM^{p-1}A \le B^p - pM^{p-1}B.$$

The choice $M := \|B\|$ yields (7.3.33). By the same method, the inequality (7.3.34) is obvious by (7.3.32). □

Remark 7.3.10. Theorem 7.3.8 shows that if $\|A\|\mathbf{1}_{\mathcal{H}} \le B$, then

$$0 \le \frac{p\|B\|^{p-1}}{\|(B-A)^{-1}\|}\mathbf{1}_{\mathcal{H}} \le p\|B\|^{p-1}(B - A) \le B^p - A^p, \quad 0 \le p \le 1.$$

We compare this with the following result given in [186, Corollary 2.5 (i)]:

$$0 \le \|B\|^p\mathbf{1}_{\mathcal{H}} - \left(\|B\| - \frac{1}{\|(B-A)^{-1}\|}\right)^p \mathbf{1}_{\mathcal{H}} \le B^p - A^p,$$

for $0 \le p \le 1$ and $0 \le A < B$. Since $1 \le \|B(B-A)^{-1}\| \le \|B\|\|(B-A)^{-1}\| =: s$, we show

$$ps^{p-1} \le s^p - (s-1)^p, \quad 0 \le p \le 1, s \ge 1. \tag{7.3.35}$$

Putting $t = \frac{1}{s}$, the inequality (7.3.35) is equivalent to the inequality

$$(1-t)^p \le 1 - pt, \quad 0 \le p \le 1, 0 < t \le 1.$$

This inequality can be proven by putting $f_p(t) = 1 - pt - (1-t)^p$ and then calculating $f_p'(t) = p\{(1-t)^{p-1} - 1\} \ge 0$ which implies $f_p(t) \ge f_p(0) = 0$. The inequality (7.3.35) thus implies the relation

$$\frac{p\|B\|^{p-1}}{\|(B-A)^{-1}\|}\mathbf{1}_{\mathcal{H}} \le \|B\|^p \mathbf{1}_{\mathcal{H}} - \left(\|B\| - \frac{1}{\|(B-A)^{-1}\|}\right)^p \mathbf{1}_{\mathcal{H}}.$$

We can give more applications of Lemma 7.3.1 and many hidden consequences of Lemma 7.3.1 as some inequalities in an inner product space and a norm inequality. The obtained results improve the known classical inequalities in the below.

The following is an extension of the result by B. Mond and J. Pečarić [97, Theorem 1.2] from convex functions to differentiable functions.

Theorem 7.3.9 (A weakened version of Mond–Pečarić inequality [175]). *Let $A, B \in \mathbb{B}(\mathcal{H})$ such that $0 < m\mathbf{1}_{\mathcal{H}} \le B \le A \le M\mathbf{1}_{\mathcal{H}}$. If f is a continuous differentiable function such that $\alpha \le f' \le \beta$ with $\alpha, \beta \in \mathbb{R}$, then*

$$\alpha\big\langle (A-B)x, x\big\rangle \le \big\langle f(A)x, x\big\rangle - f\big(\langle Bx, x\rangle\big) \le \beta\big\langle (A-B)x, x\big\rangle,$$

for every unit vector $x \in \mathcal{H}$.

Proof. We follow a similar path to the proof of [183, Theorem 3.3]. Due to relation (7.3.27), we have

$$a\mathbf{1}_{\mathcal{H}} \le A \Rightarrow \begin{cases} f(a)\mathbf{1}_{\mathcal{H}} - \alpha a\mathbf{1}_{\mathcal{H}} \le f(A) - \alpha A \\ \beta a\mathbf{1}_{\mathcal{H}} - f(a)\mathbf{1}_{\mathcal{H}} \le \beta A - f(A). \end{cases}$$

So for all $x \in \mathcal{H}$ with $\|x\| = 1$, we have

$$a \le \langle Ax, x\rangle \Rightarrow \begin{cases} f(a) - \alpha a \le \big\langle f(A)x, x\big\rangle - \alpha\langle Ax, x\rangle \\ \beta a - f(a) \le \beta\langle Ax, x\rangle - \big\langle f(A)x, x\big\rangle. \end{cases}$$

With the substitution $a = \langle Bx, x\rangle$, this becomes

$$\langle Bx, x\rangle \le \langle Ax, x\rangle \Rightarrow \begin{cases} f\big(\langle Bx, x\rangle\big) - \alpha\langle Bx, x\rangle \le \big\langle f(A)x, x\big\rangle - \alpha\langle Ax, x\rangle \\ \beta\langle Bx, x\rangle - f\big(\langle Bx, x\rangle\big) \le \beta\langle Ax, x\rangle - \big\langle f(A)x, x\big\rangle, \end{cases}$$

which is the desired conclusion. □

The next theorem was given in [156] in a more general setting.

Theorem 7.3.10 (Hölder–McCarthy inequality [156]). *Let $A \in \mathbb{B}(\mathcal{H})$ be a positive operator and $x \in \mathcal{H}$ be a unit vector.*

(1) $\langle A^p x, x\rangle \le \langle Ax, x\rangle^p$ *for all* $0 < p < 1$.
(2) $\langle Ax, x\rangle^p \le \langle A^p x, x\rangle$ *for all* $p > 1$.
(3) *If* A *is invertible, then* $\langle Ax, x\rangle^p \le \langle A^p x, x\rangle$ *for all* $p < 0$.

Using Lemma 7.3.1, we are able to point out that the following inequalities give a reverse and an improvement of the Hölder–McCarthy inequality.

Proposition 7.3.4. *For* $A > 0$ *with* $0 < m\mathbf{1}_{\mathcal{H}} \le A \le M\mathbf{1}_{\mathcal{H}}$ *and a unit vector* $x \in \mathcal{H}$*, we have*
(i) *If* $0 < p < 1$*, then*

$$\frac{p}{M^{1-p}}\left(\langle Ax, x\rangle - \langle A^p x, x\rangle^{\frac{1}{p}}\right) \le \langle Ax, x\rangle^p - \langle A^p x, x\rangle \le \frac{p}{m^{1-p}}\left(\langle Ax, x\rangle - \langle A^p x, x\rangle^{\frac{1}{p}}\right).$$

(ii) *If* $p > 1$ *or* $p < 0$*, then*

$$\frac{p}{m^{1-p}}\left(\langle A^p x, x\rangle^{\frac{1}{p}} - \langle Ax, x\rangle\right) \le \langle A^p x, x\rangle - \langle Ax, x\rangle^p \le \frac{p}{M^{1-p}}\left(\langle A^p x, x\rangle^{\frac{1}{p}} - \langle Ax, x\rangle\right).$$

Proof. The first one follows from (7.3.28) by taking $a = \langle A^p x, x\rangle^{\frac{1}{p}}$, $b = \langle Ax, x\rangle$ with $0 < p < 1$. The second one is completely similar as that before, so we leave out the details. □

Let $A \in \mathbb{B}(\mathcal{H})$. As it is well known,

$$\|A\| \le \|A\|_2 \le \|A\|_1, \tag{7.3.36}$$

where $\|\cdot\|$, $\|\cdot\|_2$, $\|\cdot\|_1$ are usual operator norm, Hilbert–Schmidt norm and trace class norm, respectively. As a consequence of the inequality (7.3.27), we have a refinement and a reverse of (7.3.36) as follows:

$$\|A\| \le \frac{\|A\|_2^p - \|A\|^p}{p\|A\|_1^{p-1}} + \|A\| \le \|A\|_2 \le \frac{\|A\|_1^p - \|A\|_2^p}{p\|A\|_1^{p-1}} + \|A\|_2 \le \|A\|_1, \quad p \ge 1 \text{ or } p \le 0$$

and

$$\frac{\|A\|_2^p - \|A\|_1^p}{p\|A\|_1^{p-1}} + \|A\|_1 \le \|A\|_2 \le \frac{\|A\|_2^p - \|A\|^p}{p\|A\|_1^{p-1}} + \|A\|, \quad 0 < p \le 1.$$

More in the same vein as above, if $A \in \mathbb{B}(\mathcal{H})$, then

$$r(A) \le \frac{w^p(A) - r^p(A)}{p\|A\|^{p-1}} + r(A) \le w(A) \le \frac{\|A\|^p - w^p(A)}{p\|A\|^{p-1}} + w(A) \le \|A\|, \quad p \ge 1 \text{ or } p \le 0$$

and

$$\frac{w^p(A) - \|A\|^p}{p\|A\|^{p-1}} + \|A\| \le w(A) \le \frac{w^p(A) - r^p(A)}{p\|A\|^{p-1}} + r(A), \quad 0 < p \le 1,$$

where $r(A)$, $w(A)$ and $\|A\|$ are the spectral radius, numerical radius and the usual operator norm of A, respectively. The inequalities above follow from the fact that for any $A \in \mathbb{B}(\mathcal{H})$,

$$r(A) \leq w(A) \leq \|A\|.$$

The following norm inequalities are well known.

Theorem 7.3.11 ([20, Theorem IX.2.1 and IX.2.3]). *For $A, B > 0$, we have:*
(i) *If $0 \leq p \leq 1$, then*

$$\|A^p B^p\| \leq \|AB\|^p. \tag{7.3.37}$$

(ii) *If $p \geq 1$, then*

$$\|AB\|^p \leq \|A^p B^p\|. \tag{7.3.38}$$

The following proposition provides a refinement and a reverse for the inequalities (7.3.37) and (7.3.38). The proof is the same as that of Proposition 7.3.2 and we omit it.

Proposition 7.3.5. *Let A, B be two positive operators such that $0 < m\mathbf{1}_{\mathcal{H}} \leq A, B \leq M\mathbf{1}_{\mathcal{H}}$. Then*

$$\frac{p}{M^{1-p}}\left(\|AB\| - \|A^p B^p\|^{\frac{1}{p}}\right) \leq \|AB\|^p - \|A^p B^p\| \leq \frac{p}{m^{1-p}}\left(\|AB\| - \|A^p B^p\|^{\frac{1}{p}}\right), \tag{7.3.39}$$

for any $0 \leq p \leq 1$. Moreover, if $p \geq 1$, then

$$\frac{p}{m^{1-p}}\left(\|A^p B^p\|^{\frac{1}{p}} - \|AB\|\right) \leq \|A^p B^p\| - \|AB\|^p \leq \frac{p}{M^{1-p}}\left(\|A^p B^p\|^{\frac{1}{p}} - \|AB\|\right). \tag{7.3.40}$$

Remark 7.3.11. One can construct other norm (trace and determinant) inequalities using our approach given in Lemma 7.3.1. We leave the details of this idea to the interested reader, as it is just an application of our main results in this section.

8 Miscellaneous topics

8.1 Kantorovich inequality

In [124], L. V. Kantorovich, Soviet mathematician and economist, introduced the well-known Kantorovich inequality. Operator version of Kantorovich inequality was first established by A. W. Marshall and I. Olkin, who obtained the following.

Theorem 8.1.1 ([153]). *Let A be a positive operator satisfying $0 < mI \le A \le MI$ for some scalars m, M with $m < M$ and Φ be a normalized positive linear map. Then*

$$\Phi(A^{-1}) \le \frac{(M+m)^2}{4Mm}\Phi^{-1}(A). \tag{8.1.1}$$

8.1.1 New operator Kantorovich inequalities

We aim to present an improvement of the inequality (8.1.1). The main result of this subsection is of this genre.

Theorem 8.1.2. *Under same assumptions of Theorem* 8.1.1*, we have*

$$\Phi(A^{-1}) \le \Phi\left(m^{\frac{A-MI}{M-m}} M^{\frac{mI-A}{M-m}}\right) \le \frac{(M+m)^2}{4Mm}\Phi^{-1}(A). \tag{8.1.2}$$

This will be proven in the below.

A positive function defined on the interval J (or, more generally, on a convex subset of some vector space) is called **log-convex** if $\log f(x)$ is a convex function of $x \in J$. We observe that such functions satisfy the elementary inequality

$$f((1-v)a + vb) \le f^{1-v}(a) f^{v}(b), \quad v \in [0,1]$$

for any $a, b \in J$. Because of the weighted arithmetic–geometric mean inequality, we have

$$f((1-v)a + vb) \le f^{1-v}(a) f^{v}(b) \le (1-v)f(a) + vf(b), \tag{8.1.3}$$

which states that all log-convex function is a convex function. The following inequality is well known in the literature as the C-D-J inequality.

Theorem 8.1.3 ([34, 38]). *Let $A \in \mathbb{B}(\mathcal{H})$ be a self-adjoint operator with spectrum $Sp(A) \subseteq J$ and Φ be a normalized positive linear map from $\mathbb{B}(\mathcal{H})$ to $\mathbb{B}(\mathcal{K})$. If f is an operator convex function on an interval J, then*

$$f(\Phi(A)) \le \Phi(f(A)). \tag{8.1.4}$$

https://doi.org/10.1515/9783110643473-008

Though in the case of convex function the inequality (8.1.4) does not hold in general, we have the following estimate.

Theorem 8.1.4 ([164, Remark 4.14]). *Let $A \in \mathbb{B}(\mathcal{H})$ be a self-adjoint operator with $Sp(A) \subseteq [m, M]$ for some scalars m, M with $m < M$ and $\Phi : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ be a normalized positive linear map. If f is a nonnegative convex function, then*

$$\frac{1}{K(m,M,f)}\Phi(f(A)) \le f(\Phi(A)) \le K(m,M,f)\Phi(f(A)),$$

where $K(m,M,f)$ is defined in (5.2.11).

Below we prove an analogue of Theorem 8.1.4 for log-convex functions. The proof of Theorem 8.1.2 follows quickly from this inequality. Inspired by the work of M. Lin [143], we square the second inequality in (8.1.2). An important role in our analysis is played by the following result, which is of independent interest.

Proposition 8.1.1. *Under same assumptions of Theorem* 8.1.4, *except that the condition convexity is changed to* log-*convexity, we have*

$$\Phi(f(A)) \le \Phi(f^{\frac{MI-A}{M-m}}(m) f^{\frac{A-mI}{M-m}}(M)) \le K(m,M,f) f(\Phi(A)). \tag{8.1.5}$$

Proof. It can be verified that if $m \le t \le M$, then $0 \le \frac{M-t}{M-m}, \frac{t-m}{M-m} \le 1$ and $\frac{M-t}{M-m} + \frac{t-m}{M-m} = 1$. Thanks to (8.1.3), we have

$$f(t) = f\left(\frac{M-t}{M-m}m + \frac{t-m}{M-m}M\right) \le f^{\frac{M-t}{M-m}}(m) f^{\frac{t-m}{M-m}}(M) \le L(t), \tag{8.1.6}$$

where $L(t)$ is defined in (5.2.8). Applying functional calculus for the operator A, we infer that

$$f(A) \le f^{\frac{MI-A}{M-m}}(m) f^{\frac{A-mI}{M-m}}(M) \le L(A).$$

Using the hypotheses on Φ,

$$\Phi(f(A)) \le \Phi(f^{\frac{MI-A}{M-m}}(m) f^{\frac{A-mI}{M-m}}(M)) \le \Phi(L(A)). \tag{8.1.7}$$

On account of [164, Corollary 4.12], we get

$$\Phi(f(A)) \le \Phi(f^{\frac{MI-A}{M-m}}(m) f^{\frac{A-mI}{M-m}}(M)) \le K(m,M,f) f(\Phi(A)).$$

Notice that, although it is known to be true for matrices [164, Corollary 4.12], it is also true for operators. Hence (8.1.5) follows. □

The following corollary follows immediately from Proposition 8.1.1. Recall that $f(t) = t^p$, $(p < 0)$ is log-convex function.

Corollary 8.1.1. *Under the hypotheses of Proposition* 8.1.1, *for* $p \in (-\infty, 0)$ *and* $0 < m < M$,

$$\Phi(A^p) \le \Phi(m^{p(\frac{MI-A}{M-m})} M^{p(\frac{A-mI}{M-m})}) \le K(m, M, p)\Phi^p(A), \tag{8.1.8}$$

where $K(m, M, p)$ *is the generalized Kantorovich constant defined in* (2.0.7).

After the previous technical intermission, we return to the main subject of this subsection, the proof of the inequality (8.1.2).

Proof of Theorem 8.1.2. This follows from Corollary 8.1.1 by putting $p = -1$. We should point out that $K(m, M, -1) = \frac{(M+m)^2}{4Mm}$. □

Can the second inequality in (8.1.2) be squared? Responding to this question is the main motivation of the next topic. We will need the following lemma.

Lemma 8.1.1. *For each* $m \le t \le M$, *we have* $t + mMm^{\frac{t-M}{M-m}} M^{\frac{m-t}{M-m}} \le M + m$.

Proof. By the weighted arithmetic–geometric mean inequality, we have

$$t + mMm^{\frac{t-M}{M-m}} M^{\frac{m-t}{M-m}} = t + m^{\frac{t-m}{M-m}} M^{\frac{M-t}{M-m}} \le t + \frac{t-m}{M-m} m + \frac{M-t}{M-m} M = M + m. \qquad \square$$

We give our main result.

Theorem 8.1.5 ([180]). *Under same assumptions of Theorem* 8.1.1, *we have*

$$\Phi^p(m^{\frac{A-MI}{M-m}} M^{\frac{mI-A}{M-m}}) \le \left(\frac{(M+m)^2}{4^{\frac{2}{p}} Mm}\right)^p \Phi^{-p}(A), \quad (2 \le p < \infty). \tag{8.1.9}$$

In particular,

$$\Phi^2(m^{\frac{A-MI}{M-m}} M^{\frac{mI-A}{M-m}}) \le \left(\frac{(M+m)^2}{4Mm}\right)^2 \Phi^{-2}(A).$$

Proof. From Lemma 3.4.1, we have only to prove

$$A \le \alpha B \Leftrightarrow \|A^{\frac{1}{2}} B^{-\frac{1}{2}}\| \le \alpha^{\frac{1}{2}}.$$

So the proof will be done if we can show

$$\|\Phi^{\frac{p}{2}}(m^{\frac{A-MI}{M-m}} M^{\frac{mI-A}{M-m}})\Phi^{\frac{p}{2}}(A)\| \le \frac{(M+m)^p}{4M^{\frac{p}{2}} m^{\frac{p}{2}}}.$$

On account of Lemma 8.1.1, it follows that

$$\Phi(A) + mM\Phi(m^{\frac{A-MI}{M-m}} M^{\frac{mI-A}{M-m}}) \le (M+m)I. \tag{8.1.10}$$

By direct calculations,

$$\begin{aligned}
&\left\|m^{\frac{p}{2}}M^{\frac{p}{2}}\Phi^{\frac{p}{2}}(m^{\frac{A-MI}{M-m}}M^{\frac{mI-A}{M-m}})\Phi^{\frac{p}{2}}(A)\right\| \\
&\le \frac{1}{4}\left\|m^{\frac{p}{2}}M^{\frac{p}{2}}\Phi^{\frac{p}{2}}(m^{\frac{A-MI}{M-m}}M^{\frac{mI-A}{M-m}}) + \Phi^{\frac{p}{2}}(A)\right\|^2 \quad \text{(by Lemma 2.6.2)} \\
&\le \frac{1}{4}\left\|\left(mM\Phi(m^{\frac{A-MI}{M-m}}M^{\frac{mI-A}{M-m}}) + \Phi(A)\right)^{\frac{p}{2}}\right\|^2 \quad \text{(by Lemma 3.4.4)} \\
&= \frac{1}{4}\left\|mM\Phi(m^{\frac{A-MI}{M-m}}M^{\frac{mI-A}{M-m}}) + \Phi(A)\right\|^p \le \frac{(M+m)^p}{4} \quad \text{(by (8.1.10))}.
\end{aligned}$$

This proves (8.1.9). □

The *Löwner–Heinz inequality* asserts that $0 \le A \le B$ ensures $A^p \le B^p$ for any $p \in [0,1]$; as it is well known that the Löwner–Heinz inequality does not always hold for $p > 1$. The following theorem is due to Furuta [91].

Theorem 8.1.6 ([91, Theorem 2.1]). *For $A, B > 0$ such that $A \le B$ and $mI \le A \le MI$ for some scalars $0 < m < M$, we have*

$$A^p \le K(m,M,p)B^p \le \left(\frac{M}{m}\right)^{p-1} B^p \quad \text{for } p \ge 1,$$

where $K(m,M,p)$ is a generalized Kantorovich constant.

In the paper [163], J. Mićić, J. Pečarić and Y. Seo proved some fascinating results about the function preserving the operator order, under a general setting.

Theorem 8.1.7 ([163, Theorem 2.1]). *Let A and B be two strictly positive operators satisfying $mI \le A \le MI$ for some scalars $0 < m < M$. Let $f : [m,M] \to \mathbb{R}$ be a convex function and $g : J \to \mathbb{R}$, where J be any interval containing $Sp(B) \cup [m,M]$. Suppose that either of the following conditions holds: (i) g is increasing convex on J, or (ii) g is decreasing concave on J. If $A \le B$, then for a given $\alpha > 0$ in the case (i) or $\alpha < 0$ in the case (ii)*

$$f(A) \le \alpha g(B) + \beta I,$$

holds for

$$\beta = \max_{m \le t \le M}\{a_f t + b_f - \alpha g(t)\}, \tag{8.1.11}$$

where

$$a_f = \frac{f(M) - f(m)}{M-m} \quad \text{and} \quad b_f = \frac{Mf(m) - mf(M)}{M-m}.$$

The following inequality is a converse of Theorem 8.1.7.

Theorem 8.1.8 ([200, Theorem 2.1]). *Let A and B be two strictly positive operators on a Hilbert space $\mathcal{H}$ satisfying $mI \le B \le MI$ for some scalars $0 < m < M$. Let $f : [m, M] \to \mathbb{R}$ be a convex function and $g : J \to \mathbb{R}$, where J be any interval containing $Sp(A) \cup [m, M]$. Suppose that either of the following conditions holds: (i) g is decreasing convex on J, or (ii) g is increasing concave on J. If $A \le B$, then for a given $\alpha > 0$ in the case (i) or $\alpha < 0$ in the case (ii)*

$$f(B) \le \alpha g(A) + \beta I, \tag{8.1.12}$$

holds with β as (8.1.11).

8.1.2 Functions reversing operator order

Our principal result is the following theorem.

Theorem 8.1.9 ([82]). *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two self-adjoint operators such that $mI \le B \le MI$ for some scalars $m < M$. Let $f : [m, M] \to (0, \infty)$ be a log-convex function and $g : J \to \mathbb{R}$, where J be any interval containing $Sp(A) \cup [m, M]$. Suppose that either of the following conditions holds: (i) g is decreasing convex on J, or (ii) g is increasing concave on J. If $A \le B$, then for a given $\alpha > 0$ in the case (i) or $\alpha < 0$ in the case (ii)*

$$f(B) \le \exp\left(\frac{MI - B}{M - m} \ln f(m) + \frac{B - mI}{M - m} \ln f(M)\right) \le \alpha g(A) + \beta I, \tag{8.1.13}$$

holds with β as (8.1.11).

Proof. We prove the inequalities (8.1.13) under the assumption (i). It is immediate to see that

$$f(t) \le f^{\frac{M-t}{M-m}}(m) f^{\frac{t-m}{M-m}}(M) \le L(t) \quad \text{for } m \le t \le M, \tag{8.1.14}$$

where $L = a_f t + b_f$ is defined in (5.2.11) (see also Corollary 7.2.1.) By applying the standard functional calculus of the self-adjoint operator B to (8.1.14), we obtain for each unit vector $x \in \mathcal{H}$,

$$\langle f(B)x, x\rangle \le \left\langle \exp\left(\frac{MI - B}{M - m} \ln f(m) + \frac{B - mI}{M - m} \ln f(M)\right)x, x\right\rangle \le a_f \langle Bx, x\rangle + b_f,$$

and from this, it follows that

$$\begin{aligned}
&\langle f(B)x, x\rangle - \alpha g(\langle Ax, x\rangle) \\
&\quad \le \left\langle \exp\left(\frac{MI - B}{M - m} \ln f(m) + \frac{B - mI}{M - m} \ln f(M)\right)x, x\right\rangle - \alpha g(\langle Ax, x\rangle) \\
&\quad \le a_f \langle Bx, x\rangle + b_f - \alpha g(\langle Ax, x\rangle) \le \max_{m \le t \le M} \{a_f t + b_f - \alpha g(t)\}.
\end{aligned}$$

Where

$$
\begin{aligned}
\langle f(B)x,x\rangle \le \left\langle \exp\left(\frac{MI-B}{M-m}\ln f(m)+\frac{B-mI}{M-m}\ln f(M)\right)x,x\right\rangle &\le \alpha g(\langle Bx,x\rangle)+\beta \\
&\le \alpha g(\langle Ax,x\rangle)+\beta \quad (\text{since } A\le B \text{ and } g \text{ is decreasing}) \\
&\le \alpha\langle g(A)x,x\rangle+\beta \quad (\text{since } g \text{ is convex}),
\end{aligned}
$$

and the assertion follows. □

The following corollary improves the result in [200, Corollary 2.5]. In fact, if we put $f(t)=t^p$ and $g(t)=t^q$ with $p\le 0$ and $q\le 0$, then we get the following.

Corollary 8.1.2. *For $A,B>0$, such that $A\le B$, $mI\le B\le MI$ for some scalars $0<m<M$ and for a given $\alpha>0$, we have*

$$
B^p \le \exp\left(\frac{MI-B}{M-m}\ln m^p+\frac{B-mI}{M-m}\ln M^p\right)\le \alpha A^q+\beta I, \quad (p,q\le 0), \tag{8.1.15}
$$

where β is defined as

$$
\beta=\begin{cases}\alpha(q-1)\left(\frac{M^p-m^p}{\alpha q(M-m)}\right)^{\frac{q}{q-1}}+\frac{Mm^p-mM^p}{M-m} & \text{if } m\le\left(\frac{M^p-m^p}{\alpha q(M-m)}\right)^{\frac{1}{q-1}}\le M\\ \max\{m^p-\alpha m^q, M^p-\alpha M^q\} & \text{otherwise.}\end{cases} \tag{8.1.16}
$$

Especially, by setting $p=q$ in (8.1.15), we reach

$$
B^p \le \exp\left(\frac{MI-B}{M-m}\ln m^p+\frac{B-mI}{M-m}\ln M^p\right)\le \alpha A^p+\beta I \quad \text{for } p\le 0,
$$

where

$$
\beta=\begin{cases}\alpha(p-1)\left(\frac{M^p-m^p}{\alpha p(M-m)}\right)^{\frac{p}{p-1}}+\frac{Mm^p-mM^p}{M-m} & \text{if } m\le\left(\frac{M^p-m^p}{\alpha p(M-m)}\right)^{\frac{1}{p-1}}\le M\\ \max\{m^p-\alpha m^p, M^p-\alpha M^p\} & \text{otherwise.}\end{cases}
$$

If we choose α such that $\beta=0$ in Theorem 8.1.9, then we obtain the following corollary. For completeness, we sketch the proof.

Corollary 8.1.3. *Let $A,B\in\mathbb{B}(\mathcal{H})$ be two strictly positive operators such that $A\le B$ and $mI\le B\le MI$ for some scalars $0<m<M$. Let $f:[m,M]\to(0,\infty)$ be a log-convex function and $g:J\to\mathbb{R}$ be a continuous function, where J is an interval containing $Sp(A)\cup[m,M]$. If g is a nonnegative decreasing convex on $[m,M]$, then*

$$
f(B)\le \exp\left(\frac{MI-B}{M-m}\ln f(m)+\frac{B-mI}{M-m}\ln f(M)\right)\le \max_{m\le t\le M}\left\{\frac{a_f t+b_f}{g(t)}\right\}g(A). \tag{8.1.17}
$$

If we take $f(t)=t^p$ and $g(t)=t^q$ with $p\le 0$ and $-1\le q\le 0$ in (8.1.17), we have

$$
B^p \le \exp\left(\frac{MI-B}{M-m}\ln m^p+\frac{B-mI}{M-m}\ln M^p\right)\le K(m,M,p,q)A^q, \tag{8.1.18}
$$

where $K(m,M,p,q)$ is defined in

$$K(m,M,p,q)=\begin{cases}\frac{(mM^p-Mm^p)}{(q-1)(M-m)}\left(\frac{q-1}{q}\frac{M^p-m^p}{mM^p-Mm^p}\right)^q & \text{if } m\le\frac{q(mM^p-Mm^p)}{(q-1)(M^p-m^p)}\le M\\ \max\{m^{p-q},M^{p-q}\} & \text{otherwise.}\end{cases}\tag{8.1.19}$$

We mention that $K(m,M,p,q)$ was given in [163, *Theorem 3.1*].

In particular, if $p=q$ in (8.1.18), *we get*

$$B^p\le\exp\left(\frac{MI-B}{M-m}\ln m^p+\frac{B-mI}{M-m}\ln M^p\right)\le K(m,M,p)A^p\quad \text{for } p\le 0,\tag{8.1.20}$$

where $K(m,M,p)$ is defined in (2.0.7).

Proof. From the condition on the function g, we have $\beta\le\max\limits_{m\le t\le M}\{a_ft+b_f\}-\alpha\min\limits_{m\le t\le M}\{g(t)\}$. When $\beta=0$, we have $\alpha\le\frac{\max_{m\le t\le M}\{a_ft+b_f\}}{\min_{m\le t\le M}\{g(t)\}}$. Thus we have the inequalities (8.1.17), taking $\alpha:=\max\limits_{m\le t\le M}\{\frac{a_ft+b_f}{g(t)}\}$ which is less than or equal to $\frac{\max_{m\le t\le M}\{a_ft+b_f\}}{\min_{m\le t\le M}\{g(t)\}}$.

If we take $f(t)=t^p$ and $g(t)=t^q$ with $p\le 0$ and $-1\le q\le 0$ for $t>0$ in (8.1.17), then we have $a_{t^p}\le 0$, $b_{t^p}\ge 0$ and $\alpha=\max\limits_{m\le t\le M}\{a_{t^p}t^{1-q}+b_{t^p}t^{-q}\}$. Then we set $h_{p,q}(t):=a_{t^p}t^{1-q}+b_{t^p}t^{-q}$. We easily calculate

$$h'_{p,q}(t)=t^{-q-1}\{(1-q)a_{t^p}t-qb_{t^p}\},\quad h''_{p,q}(t)=t^{-q-2}\{q(q-1)a_{t^p}t+q(q+1)b_{t^p}\}\le 0.$$

We find $h_{p,q}(t)$ is concave in t and $\alpha=\max\limits_{m\le t\le M}h_{p,q}(t)=h_{p,q}(t_0)=(\frac{b_{t^p}}{1-q})\{\frac{(1-q)a_{t^p}}{qb_{t^p}}\}^q$ if $t_0=\frac{qb_{t^p}}{(1-q)a_{t^p}}$ satisfies $m\le t_0\le M$. Thus we have $\alpha=K(m,M,p,q)$ by simple calculations with $a_{t^p}=(M^p-m^p)/(M-m)$, $b_{t^p}=(Mm^p-mM^p)/(M-m)$ and the other cases are trivial. Thus we have the inequalities (8.1.18) and (8.1.20). □

This corollary is a refinement of [200, Corollary 2.6]. We give a refinement of [126, Corollary 2.2] (see also [161, Corollary 1]) as follows.

Corollary 8.1.4. *For $A,B>0$ such that $A\le B$ and $mI\le B\le MI$ for some scalars $0<m<M$, we have*

$$B^p\le\exp\left(\frac{MI-B}{M-m}\ln m^p+\frac{B-mI}{M-m}\ln M^p\right)\le C(m,M,p,q)I+A^q\quad \text{for } p,q\le 0,\tag{8.1.21}$$

where $C(m,M,p,q)$ is the Kantorovich constant for the difference with two parameters and defined by

$$C(m,M,p,q)=\begin{cases}\frac{Mm^p-mM^p}{M-m}+(q-1)\left(\frac{M^p-m^p}{q(M-m)}\right)^{\frac{q}{q-1}} & \text{if } m\le\left(\frac{M^p-m^p}{q(M-m)}\right)^{\frac{1}{q-1}}\le M\\ \max\{M^p-M^q,m^p-m^q\} & \text{otherwise.}\end{cases}$$

Proof. If we put $\alpha = 1$, $f(t) = t^p$ for $p \le 0$ and $g(t) = t^q$ for $q \le 0$ in Theorem 8.1.9, then we have $\beta = \max_{m \le t \le M} \{a_{t^p} t + b_{t^p} - t^q\}$. By simple calculations, we have $\beta = (q-1)(\frac{a_{t^p}}{q})^{\frac{q}{q-1}} + b_{t^p}$ when $t_0 = (\frac{a_{t^p}}{q})^{\frac{1}{q-1}}$ satisfies $m \le t_0 \le M$. The other cases are trivial. Thus we have the desired conclusion, since $a_{t^p} = (M^p - m^p)/(M-m)$ and $b_{t^p} = (Mm^p - mM^p)/(M-m)$. □

We conclude this subsection by presenting some inequalities on **chaotic order** (i. e., $\log A \le \log B$ for $A, B > 0$). To achieve our next results, we need the following lemma. Its proof is standard but we provide a proof for the sake of completeness.

Lemma 8.1.2. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two strictly positive operators. Then the following statements are equivalent:*
(i) $\log A \le \log B$.
(ii) $B^r \le (B^{\frac{r}{2}} A^p B^{\frac{r}{2}})^{\frac{r}{p+r}}$ for $p \le 0$ and $r \le 0$.

Proof. From the well-known **chaotic Furuta inequality** (see, e. g., [56, 90]) the order $\log A \ge \log B$ is equivalent to the inequality $(B^{\frac{r}{2}} A^p B^{\frac{r}{2}})^{\frac{p}{p+r}} \ge B^r$ for $p, r \ge 0$ and $A, B > 0$. The assertion (i) is equivalent to the order $\log B^{-1} \le \log A^{-1}$. By the use of chaotic Furuta inequality, the order $\log B^{-1} \le \log A^{-1}$ is equivalent to the inequality

$$B^{-r} \le (B^{\frac{-r}{2}} A^{-p} B^{\frac{-r}{2}})^{\frac{r}{p+r}} \quad \text{for } p, r \ge 0. \tag{8.1.22}$$

This is equivalent to the inequality

$$B^{r'} \le (B^{\frac{r'}{2}} A^{p'} B^{\frac{r'}{2}})^{\frac{r'}{p'+r'}} \quad \text{for } p', r' \le 0,$$

by substituting $p' = -p$ and $r' = -r$ in (8.1.22). We thus obtain the desired conclusion. □

As an application of Corollary 8.1.3, we have the following result.

Corollary 8.1.5. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two strictly positive operators such that $mI \le B \le MI$ for some scalars $0 < m < M$ and $\log A \le \log B$. Then for $p \le 0$ and $-1 \le r \le 0$,*

$$B^p \le B^{-r} \exp\left(\frac{MI - B}{M - m} \ln m^{p+r} + \frac{B - mI}{M - m} \ln M^{p+r}\right) \le K(m, M, p + r) A^p.$$

Proof. Thanks to Lemma 8.1.2, the chaotic order $\log A \le \log B$ is equivalent to $B^r \le (B^{\frac{r}{2}} A^p B^{\frac{r}{2}})^{\frac{r}{p+r}}$ for $p, r \le 0$. Putting $B_1 = B$ and $A_1 = (B^{\frac{r}{2}} A^p B^{\frac{r}{2}})^{\frac{1}{p+r}}$ in the above, then $0 < A_1 \le B_1$ and $mI \le B_1 \le MI$. Thus we have for $p_1 \le 0$,

$$\begin{aligned} B^{p_1} = B_1^{p_1} &\le \exp\left(\frac{MI - B}{M - m} \ln m^{p_1} + \frac{B - mI}{M - m} \ln M^{p_1}\right) \\ &\le K(m, M, p_1) A_1^{p_1} = K(m, M, p_1)(B^{\frac{r}{2}} A^p B^{\frac{r}{2}})^{\frac{p_1}{p+r}}, \end{aligned}$$

by (8.1.20). By setting $p_1 = p + r \le 0$ and multiplying $B^{-\frac{r}{2}}$ to both sides, we obtain the desired conclusion. □

In a similar fashion, one can prove the following result.

Corollary 8.1.6. *Let $A, B \in \mathbb{B}(\mathcal{H})$ be two strictly positive operators such that $mI \le B \le MI$ for some scalars $0 < m < M$ and $\log A \le \log B$. Then for $p \le 0$ and $-1 \le r \le 0$,*

$$B^p \le B^{-r} \exp\left(\frac{MI - B}{M - m} \ln m^{p+r} + \frac{B - mI}{M - m} \ln M^{p+r}\right) \le C(m, M, p + r)I + A^p.$$

Proof. If we set $p = q$ in Corollary 8.1.4, we have the following inequalities for $p \le 0$:

$$B^p \le \exp\left(\frac{MI - B}{M - m} \ln m^p + \frac{B - mI}{M - m} \ln M^p\right) \le C(m, M, p)I + A^p, \tag{8.1.23}$$

where

$$C(m, M, p) = \begin{cases} \frac{Mm^p - mM^p}{M-m} + (p-1)\left(\frac{M^p - m^p}{p(M-m)}\right)^{\frac{p}{p-1}} & \text{if } m \le \left(\frac{M^p - m^p}{p(M-m)}\right)^{\frac{1}{p-1}} \le M \\ 0 & \text{otherwise.} \end{cases}$$

Thanks to Lemma 8.1.2, the chaotic order $\log A \le \log B$ is equivalent to $B^r \le (B^{\frac{r}{2}} A^p B^{\frac{r}{2}})^{\frac{r}{p+r}}$ for $p, r \le 0$. Putting $B_1 = B$ and $A_1 = (B^{\frac{r}{2}} A^p B^{\frac{r}{2}})^{\frac{1}{p+r}}$ in the above, then $0 < A_1 \le B_1$ and $mI \le B_1 \le MI$. Thus we have for $p_1 \le 0$,

$$B_1^{p_1} \le \exp\left(\frac{MI - B_1}{M - m} \ln m^{p_1} + \frac{B_1 - mI}{M - m} \ln M^{p_1}\right) \le C(m, M, p_1)I + A_1^{p_1},$$

by (8.1.23). Putting $p_1 = p + r \le 0$ and multiplying $B^{-\frac{r}{2}}$ to both sides, we obtain the desired conclusion. ☐

8.1.3 New refined operator Kantorovich inequality

As a multiple operator version of the celebrated D-C-J inequality [34], Mond and Pečarić in [171, Theorem 1] proved the inequality

$$f\left(\sum_{i=1}^{n} w_i \Phi_i(A_i)\right) \le \sum_{i=1}^{n} w_i \Phi_i(f(A_i)), \tag{8.1.24}$$

for operator convex functions f defined on an interval J, where Φ_i $(i = 1, \dots, n)$ are normalized positive linear maps from $\mathbb{B}(\mathcal{H})$ to $\mathbb{B}(\mathcal{K})$, $A_1, \dots, A_n$ are self-adjoint operators with spectra in J and $w_1, \dots, w_n$ are nonnegative real numbers with $\sum_{i=1}^{n} w_i = 1$.

In a reverse direction to that of inequality (8.1.24), we have the following.

Theorem 8.1.10 (Reverse D-C-J inequality for multiple operators [82]). *Let $\Phi_i : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ be normalized positive linear maps, $A_i \in \mathbb{B}(\mathcal{H})$ be self-adjoint operators with $mI \le A_i \le MI$ for some scalars $m < M$ and w_i be positive numbers such that $\sum_{i=1}^{n} w_i = 1$. If f is a*

log-convex function and g is a continuous function on $[m, M]$, *then for a given* $\alpha \in \mathbb{R}$,

$$\begin{aligned}\sum_{i=1}^{n} w_i \Phi_i(f(A_i)) &\leq \sum_{i=1}^{n} w_i \Phi_i\left(\exp\left(\frac{MI - A_i}{M - m} \ln f(m) + \frac{A_i - mI}{M - m} \ln f(M)\right)\right) \\ &\leq \alpha g\left(\sum_{i=1}^{n} w_i \Phi_i(A_i)\right) + \beta I, \end{aligned} \tag{8.1.25}$$

holds for β *as* (8.1.11).

Proof. Thanks to (8.1.14), we get

$$f(A_i) \leq \exp\left(\frac{MI - A_i}{M - m} \ln f(m) + \frac{A_i - mI}{M - m} \ln f(M)\right) \leq a_f A_i + b_f I.$$

The hypotheses on Φ_i and w_i ensure the following:

$$\begin{aligned}\sum_{i=1}^{n} w_i \Phi_i(f(A_i)) &\leq \sum_{i=1}^{n} w_i \Phi_i\left(\exp\left(\frac{MI - A_i}{M - m} \ln f(m) + \frac{A_i - mI}{M - m} \ln f(M)\right)\right) \\ &\leq a_f \sum_{i=1}^{n} w_i \Phi_i(A_i) + b_f I.\end{aligned}$$

Using the fact that $mI \leq \sum_{i=1}^{n} w_i \Phi_i(A_i) \leq MI$, we can write

$$\begin{aligned}&\sum_{i=1}^{n} w_i \Phi_i(f(A_i)) - \alpha g\left(\sum_{i=1}^{n} w_i \Phi_i(A_i)\right) \\ &\leq \sum_{i=1}^{n} w_i \Phi_i\left(\exp\left(\frac{MI - A_i}{M - m} \ln f(m) + \frac{A_i - mI}{M - m} \ln f(M)\right)\right) - \alpha g\left(\sum_{i=1}^{n} w_i \Phi_i(A_i)\right) \\ &\leq a_f \sum_{i=1}^{n} w_i \Phi_i(A_i) + b_f I - \alpha g\left(\sum_{i=1}^{n} w_i \Phi_i(A_i)\right) \\ &\leq \max_{m \leq t \leq M} \{a_f t + b_f - \alpha g(t)\} I,\end{aligned}$$

which is, after a rearrangement, equivalent to (8.1.25) so the proof is complete. □

It is worth mentioning that Theorem 8.1.10 is stronger than what appears in [162, Theorem 2.2].

Theorem 8.1.11 ([82]). *Let* $A, B \in \mathbb{B}(\mathcal{H})$ *be two strictly positive operators such that* $mA \leq B \leq MA$ *for some scalars* $0 < m < M$. *Let* σ_f *and* σ_g *be two connections with the representing real valued functions f and g which are log-convex and continuous on the interval* $[m, M]$, *respectively. If either (i) g is convex on* $[m, M]$ *and* $\alpha > 0$, *or (ii) g is concave on* $[m, M]$ *and* $\alpha < 0$, *then*

$$\begin{aligned}A \sigma_f B &\leq A^{\frac{1}{2}} \exp\left(\frac{MI - A^{-\frac{1}{2}} B A^{-\frac{1}{2}}}{M - m} \ln f(m) + \frac{A^{-\frac{1}{2}} B A^{-\frac{1}{2}} - mI}{M - m} \ln f(M)\right) A^{\frac{1}{2}} \\ &\leq \alpha A \sigma_g B + \beta A,\end{aligned}$$

holds with β *as* (8.1.11).

Proof. We prove the assertion under the assumption (i). Using the same technique as in the proof of Theorem 8.1.9, we get

$$\begin{aligned}\langle f(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})x,x\rangle &\le \Big\langle \exp\Big(\frac{MI-A^{-\frac{1}{2}}BA^{-\frac{1}{2}}}{M-m}\ln f(m)+\frac{A^{-\frac{1}{2}}BA^{-\frac{1}{2}}-mI}{M-m}\ln f(M)\Big)\Big\rangle\\ &\le \alpha g(\langle A^{-\frac{1}{2}}BA^{-\frac{1}{2}}x,x\rangle)+\beta,\end{aligned}$$

where

$$\begin{aligned}f(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}) &\le \exp\Big(\frac{MI-A^{-\frac{1}{2}}BA^{-\frac{1}{2}}}{M-m}\ln f(m)+\frac{A^{-\frac{1}{2}}BA^{-\frac{1}{2}}-mI}{M-m}\ln f(M)\Big)\\ &\le \alpha g(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})+\beta I.\end{aligned}$$

Here, the last inequality is a consequence of the convexity of the function g. Multiply $A^{\frac{1}{2}}$ to both sides of the above inequality, then we get the desired result. □

Of course, we can study the particular cases $\beta = 0$ or $\alpha = 1$. But, in order to avoid repeating, we only give the following result.

Corollary 8.1.7. *For $A, B > 0$ such that $mA \le B \le MA$ for some scalars $0 < m < M$, we have*

$$A\natural_p B \le \exp\Big(\frac{MI-A^{-\frac{1}{2}}BA^{-\frac{1}{2}}}{M-m}\ln m^p+\frac{A^{-\frac{1}{2}}BA^{-\frac{1}{2}}-mI}{M-m}\ln M^p\Big)\le K(m,M,p,q)A\natural_q B,$$

for $p \le 0$ and $q \in \mathbb{R}\backslash(0,1)$, where $K(m,M,p,q)$ is defined as (8.1.19).

Before we present our theorem, we begin with a general observation.

The Kittaneh–Manasrah inequality says

$$a^{1-v}b^v \le (1-v)a+vb-r(\sqrt{a}-\sqrt{b})^2, \tag{8.1.26}$$

where $r = \min\{v, 1-v\}$. We here borrow properties of **geometrically convex** function. See Section 2.10 and (5.6.2) for its definition and some basic properties. Now, if f is a decreasing geometrically convex function, then we have

$$\begin{aligned}f((1-v)a+vb) &\le f(((1-v)a+vb)-r(\sqrt{a}-\sqrt{b})^2)\le f(a^{1-v}b^v)\le f^{1-v}(a)f^v(b)\\ &\le (1-v)f(a)+vf(b)-r\Big(\sqrt{f(a)}-\sqrt{f(b)}\Big)^2\le(1-v)f(a)+vf(b),\end{aligned}\tag{8.1.27}$$

where the first inequality follows from the inequality $(1-v)a+vb-r(\sqrt{a}-\sqrt{b})^2 \le (1-v)a+vb$ and the fact that f is decreasing function, in the second inequality we used (8.1.26), the third inequality is obvious by (5.6.2) and the fourth inequality again follows from (8.1.26) by interchanging a by $f(a)$ and b by $f(b)$.

Of course, each decreasing geometrically convex function is also convex. However, the converse does not hold in general.

The inequality (8.1.27) applied to $a = m$, $b = M$, $1 - v = \frac{M-t}{M-m}$ and $v = \frac{t-m}{M-m}$ gives

$$\begin{aligned} f(t) &\le f\big(t - (\sqrt{m} - \sqrt{M})^2 r(t)\big) \le f\big(m^{\frac{M-t}{M-m}} M^{\frac{t-m}{M-m}}\big) \le f^{\frac{M-t}{M-m}}(m) f^{\frac{t-m}{M-m}}(M) \\ &\le \frac{M-t}{M-m} f(m) + \frac{t-m}{M-m} f(M) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 r(t) \\ &\le \frac{M-t}{M-m} f(m) + \frac{t-m}{M-m} f(M) \end{aligned} \tag{8.1.28}$$

with $r(t) = \min\{\frac{t-m}{M-m}, \frac{M-t}{M-m}\} = \frac{1}{2} - \frac{1}{M-m}|t - \frac{M+m}{2}|$, whenever $t \in [m, M]$.

Now, we are in a position to state and prove our main results.

Theorem 8.1.12 ([177]). *Let $A \in \mathbb{B}(\mathcal{H})$ be a self-adjoint operator with $Sp(A) \subseteq [m, M]$ for some scalars m, M with $m < M$ and $\Phi : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ be a normalized positive linear map. If f is a nonnegative decreasing geometrically convex function, then*

$$\begin{aligned} \Phi\big(f\big(A - (\sqrt{m} - \sqrt{M})^2 r(A)\big)\big) &\le \Phi\big(f\big(m^{\frac{MI-A}{M-m}} M^{\frac{A-mI}{M-m}}\big)\big) \\ &\le K(m, M, f) f\big(\Phi(A)\big) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 \Phi\big(r(A)\big), \end{aligned}$$

where $r(A) = \min\{\frac{A-m}{M-m}, \frac{M-A}{M-m}\} = \frac{1}{2} - \frac{1}{M-m}|A - \frac{M+m}{2}|$ and $K(m, M, f)$ is defined in (5.2.11).

Proof. On account of the assumptions, from parts of (8.1.28), we have

$$\begin{aligned} f\big(t - (\sqrt{m} - \sqrt{M})^2 r(t)\big) &\le f\big(m^{\frac{M-t}{M-m}} M^{\frac{t-m}{M-m}}\big) \\ &\le L(t) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 r(t), \end{aligned} \tag{8.1.29}$$

where $L(t)$ is defined in (5.2.8). Note that inequality (8.1.29) holds for all $t \in [m, M]$. On the other hand, $Sp(A) \subseteq [m, M]$, which, by virtue of monotonicity principle (1.2.1) for functions, yields the series of inequalities.

$$f\big(A - (\sqrt{m} - \sqrt{M})^2 r(A)\big) \le f\big(m^{\frac{MI-A}{M-m}} M^{\frac{A-mI}{M-m}}\big) \le L(A) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 r(A).$$

It follows from the linearity of the map Φ that

$$\begin{aligned} \Phi\big(f\big(A - (\sqrt{m} - \sqrt{M})^2 r(A)\big)\big) &\le \Phi\big(f\big(m^{\frac{MI-A}{M-m}} M^{\frac{A-mI}{M-m}}\big)\big) \\ &\le \Phi\big(L(A)\big) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 \Phi\big(r(A)\big). \end{aligned}$$

Now, by using [164, Corollary 4.12] we get

$$\begin{aligned} \Phi\big(f\big(A - (\sqrt{m} - \sqrt{M})^2 r(A)\big)\big) &\le \Phi\big(f\big(m^{\frac{MI-A}{M-m}} M^{\frac{A-mI}{M-m}}\big)\big) \\ &\le \Phi\big(L(A)\big) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 \Phi\big(r(A)\big) \\ &\le \mu(m, M, f) f\big(\Phi(A)\big) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 \Phi\big(r(A)\big). \end{aligned}$$

This completes the proof. □

The following fact can be easily deduced from Theorem 8.1.12 and the basic properties of a geometrically convex function defined in (5.6.2).

Corollary 8.1.8. *Let $A \in \mathbb{B}(\mathcal{H})$ be a positive operator with $Sp(A) \subseteq [m, M]$ for some scalars m, M with $0 < m < M$ and $\Phi : \mathbb{B}(\mathcal{H}) \to \mathbb{B}(\mathcal{K})$ be a normalized positive linear map. Then for any $p < 0$,*

$$\begin{aligned}\Phi(A^p) &\le \Phi\big(\big(A - (\sqrt{m} - \sqrt{M})^2 r(A)\big)^p\big) \le \Phi\big(\big(m^{\frac{A-MI}{M-m}} M^{\frac{mI-A}{M-m}}\big)^p\big)\\ &\le K(m, M, p)\Phi^p(A) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 \Phi(r(A)),\end{aligned}$$

where $K(m, M, p)$ is the generalized Kantorovich constant given in (2.0.7). *In particular,*

$$\begin{aligned}\Phi(A^{-1}) &\le \Phi\big(\big(A - (\sqrt{m} - \sqrt{M})^2 r(A)\big)^{-1}\big) \le \Phi\big(\big(m^{\frac{A-MI}{M-m}} M^{\frac{mI-A}{M-m}}\big)^{-1}\big)\\ &\le \frac{(M+m)^2}{4Mm}\Phi^{-1}(A) - \left(\frac{(\sqrt{M} - \sqrt{m})^2}{Mm}\right)\Phi(r(A)).\end{aligned}$$

We note that $K(m, M, -1) = K(m, M, 2) = \frac{(M+m)^2}{4Mm}$ is the original Kantorovich constant.

Theorem 8.1.13 ([177]). *Under same assumptions of Theorem* 8.1.12*, we have*

$$\begin{aligned}f\big(\Phi(A) - (\sqrt{m} - \sqrt{M})^2 r(\Phi(A))\big) &\le f\big(m^{\frac{MI-\Phi(A)}{M-m}} M^{\frac{\Phi(A)-mI}{M-m}}\big)\\ &\le K(m, M, f)\Phi(f(A)) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 r(\Phi(A)).\end{aligned}$$

Proof. By applying a standard functional calculus for the operator $\Phi(A)$ such that $mI \le \Phi(A) \le MI$, we get from (8.1.29)

$$\begin{aligned}f\big(\Phi(A) - (\sqrt{m} - \sqrt{M})^2 r(\Phi(A))\big) &\le f\big(m^{\frac{MI-\Phi(A)}{M-m}} M^{\frac{\Phi(A)-mI}{M-m}}\big)\\ &\le \Phi(L(A)) - r\big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 r(\Phi(A)).\end{aligned}$$

We thus have

$$\begin{aligned}&f\big(\Phi(A) - (\sqrt{m} - \sqrt{M})^2 r(\Phi(A))\big) \le f\big(m^{\frac{MI-\Phi(A)}{M-m}} M^{\frac{\Phi(A)-mI}{M-m}}\big)\\ &\quad\le L(\Phi(A)) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 r(\Phi(A)) = \Phi(L(A)) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 r(\Phi(A))\\ &\quad\le \mu(m, M, f)\Phi(f(A)) - \big(\sqrt{f(m)} - \sqrt{f(M)}\big)^2 r(\Phi(A)),\end{aligned}$$

where at the last step we used the basic inequality [164, Corollary 4.12]. Hence, the proof is complete. □

As a corollary of Theorem 8.1.13, we have the following.

Corollary 8.1.9. *Under same assumptions of Corollary* 8.1.8*, we have for any $p < 0$,*

$$\begin{aligned}\Phi^p(A) &\le \big(\Phi(A) - (\sqrt{m} - \sqrt{M})^2 r(\Phi(A))\big)^p \le \big(m^{\frac{MI-\Phi(A)}{M-m}} M^{\frac{\Phi(A)-mI}{M-m}}\big)^p\\ &\le K(m, M, p)\Phi(A^p) - (\sqrt{m^p} - \sqrt{M^p})^2 r(\Phi(A)).\end{aligned}$$

8.2 Skew information and uncertainty relation

We studied some properties, relations and inequalities on several entropies in Chapter 7. Especially, quantum mechanical entropy and relative entropy are fundamental concepts in quantum information theory which attracts many researchers including mathematicians, because mathematical tools such as matrix analysis, operator theory and some techniques on inequalities are applied to obtain the advanced results. Among topics on quantum information theory, we pick up the uncertainty relations which are described by trace inequalities.

In addition, as one of the mathematical studies on entropy, skew entropy [233, 234] and its concavity problem are famous. The concavity problem for skew entropy generalized by F. J. Dyson, was proven by E. H. Lieb in [141]. It is also known that skew entropy presents the degree of noncommutativity between a certain quantum state represented by a density operator and an observable represented by a self-adjoint operator. In such situations, S. Luo and Q. Zhang studied the relation between skew information (= opposite signed skew entropy) and the uncertainty relation in [150]. Inspired by their interesting work, we define a generalized skew information and then study the relation between it and the uncertainty relation, in the following subsection.

For density operator (quantum state) ρ and self-adjoint operators (observables) A and B acting on Hilbert space $\mathcal{H}$, we denote $\widetilde{A} = A - \mathrm{Tr}[\rho A]I$ and $\widetilde{B} = B - \mathrm{Tr}[\rho B]I$, where I represents the identity operator. Then we define the **covariance** by $\mathrm{Cov}_\rho(A, B) = \mathrm{Tr}[\rho\widetilde{A}\widetilde{B}]$. Each **variance** was defined by $V_\rho(A) = \mathrm{Cov}_\rho(A, A)$ and $V_\rho(B) = \mathrm{Cov}_\rho(B, B)$.

The famous **Heisenberg's uncertainty relation** [108, 206] can be easily proven by the application of Schwarz inequality and it was given by

$$V_\rho(A)V_\rho(B) \geq \frac{1}{4}\left|\mathrm{Tr}[\rho[A, B]]\right|^2 \tag{8.2.1}$$

for a quantum state ρ and two observables A and B. The further strong result was given by Schrödinger [218] in the following.

Proposition 8.2.1 (Schrödinger uncertainty relation [218]). *For any density operator ρ and any two self-adjoint operators A and B, we have the uncertainty relation:*

$$V_\rho(A)V_\rho(B) - \left|\mathrm{Re}\{\mathrm{Cov}_\rho(A, B)\}\right|^2 \geq \frac{1}{4}\left|\mathrm{Tr}[\rho[A, B]]\right|^2, \tag{8.2.2}$$

where $[X, Y] = XY - YX$ is **commutator.**

The **Wigner–Yanase skew information**

$$I_\rho(H) = \frac{1}{2}\mathrm{Tr}[(i[\rho^{1/2}, H])^2] = \mathrm{Tr}[\rho H^2] - \mathrm{Tr}[\rho^{1/2}H\rho^{1/2}H]$$

was defined in [233]. This quantity can be considered as a kind of the degree for noncommutativity between a quantum state ρ and an observable H. This quantity was

generalized by Dyson with $\alpha \in [0,1]$:

$$I_{\rho,\alpha}(H) = \frac{1}{2}\operatorname{Tr}[(i[\rho^{\alpha}, H])(i[\rho^{1-\alpha}, H])] = \operatorname{Tr}[\rho H^2] - \operatorname{Tr}[\rho^{\alpha} H \rho^{1-\alpha} H],$$

which is known as the **Wigner–Yanase–Dyson skew information**.

To give an interpretation of skew information as a measure of quantum uncertainty and a further development of the uncertainty relation, in [150] S. Luo and Q. Zhang claimed the following uncertainty relation:

$$I_\rho(A) I_\rho(B) - \left|\operatorname{Re}\{\operatorname{Corr}_\rho(A,B)\}\right|^2 \geq \frac{1}{4}\left|\operatorname{Tr}[\rho[A,B]]\right|^2, \tag{8.2.3}$$

for two self-adjoint operators A and B, and a density operator ρ, where the correlation measure was defined by $\operatorname{Corr}_\rho(X,Y) = \operatorname{Tr}[\rho X^* Y] - \operatorname{Tr}[\rho^{1/2} X^* \rho^{1/2} Y]$ for any operator X and Y. However, we show the inequality (8.2.3) does not hold in general. We give a counterexample for (8.2.3) in Remark 8.2.2 of Subsection 8.2.1

8.2.1 Skew information and uncertainty relation

We review fundamental and useful tools to prove our results.

Lemma 8.2.1 (Cauchy–Schwarz inequality). *Let $\phi : \mathcal{L}(\mathcal{H}) \to \mathbb{C}$ be symmetric sesquilinear functional from the set of all (possibly unbounded) linear operators $\mathcal{L}(\mathcal{H})$ to complex number field $\mathbb{C}$. Then we have*

$$|\phi(X^*Y)|^2 \leq \phi(X^*X)\phi(Y^*Y),$$

for any $X, Y \in \mathcal{L}(\mathcal{H})$, if $\phi(X^*X) \geq 0$ for any $X \in \mathcal{L}(\mathcal{H})$.

We note that the following inequality holds:

$$\left|\operatorname{Re}\{\phi(AB)\}\right|^2 \leq \phi(A^2)\phi(B^2),$$

for self-adjoint operators $A, B \in \mathcal{L}(\mathcal{H})$, if we have $\phi(A^2) \geq 0$ for $A = A^* \in \mathcal{L}(\mathcal{H})$.

Applying the Bourin–Fujii inequality in Theorem 4.2.8, we have the following lemma.

Lemma 8.2.2. *For any functions f and g defined on the domain $D \subset \mathbb{R}$, any self-adjoint operators A and B, and any linear operator $X \in \mathcal{L}(\mathcal{H})$, we have*

(i) *If $\{f(a) - f(b)\}\{g(a) - g(b)\} \geq 0$ for $a, b \in D$, then*

$$\operatorname{Tr}[f(A)X^*g(B)X + f(B)Xg(A)X^*] \leq \operatorname{Tr}[f(A)g(A)X^*X + f(B)g(B)XX^*].$$

(ii) *If $\{f(a) - f(b)\}\{g(a) - g(b)\} \leq 0$ for $a, b \in D$, then*

$$\operatorname{Tr}[f(A)X^*g(B)X + f(B)Xg(A)X^*] \geq \operatorname{Tr}[f(A)g(A)X^*X + f(B)g(B)XX^*].$$

Proof. We set two self-adjoint operators on $\mathcal{H} \oplus \mathcal{H}$ such as

$$\widehat{A} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \quad \widehat{X} = \begin{pmatrix} 0 & X^* \\ X & 0 \end{pmatrix}.$$

We also set (f, g) on $\mathbb{R} \oplus \mathbb{R}$ such as $\{f(a) - f(b)\}\{g(a) - g(b)\} \geq 0$. By Theorem 4.2.8, we have

$$\begin{aligned} &\mathrm{Tr}\left[\begin{pmatrix} f(A) & 0 \\ 0 & f(B) \end{pmatrix}\begin{pmatrix} 0 & X^* \\ X & 0 \end{pmatrix}\begin{pmatrix} g(A) & 0 \\ 0 & g(B) \end{pmatrix}\begin{pmatrix} 0 & X^* \\ X & 0 \end{pmatrix}\right] \\ &\quad \leq \mathrm{Tr}\left[\begin{pmatrix} f(A) & 0 \\ 0 & f(B) \end{pmatrix}\begin{pmatrix} g(A) & 0 \\ 0 & g(B) \end{pmatrix}\begin{pmatrix} 0 & X^* \\ X & 0 \end{pmatrix}\begin{pmatrix} 0 & X^* \\ X & 0 \end{pmatrix}\right], \end{aligned}$$

which implies (i) of this lemma. (ii) is proven by the similar way. □

We define the generalized skew correlation and the generalized skew information as follows.

Definition 8.2.1. For any two self-adjoint operators A and B, any density operator ρ, any $0 < \alpha < 1$ and any $\varepsilon \geq 0$, the generalized skew correlation is defined by

$$\mathrm{Corr}_{\alpha,\rho,\varepsilon}(A, B) = \left(\frac{1}{2} + \varepsilon\right)\mathrm{Tr}[\rho\widetilde{A}\widetilde{B}] + \frac{1}{2}\mathrm{Tr}[\rho\widetilde{B}\widetilde{A}] - \frac{1}{2}\mathrm{Tr}[\rho^\alpha\widetilde{A}\rho^{1-\alpha}\widetilde{B}] - \frac{1}{2}\mathrm{Tr}[\rho^\alpha\widetilde{B}\rho^{1-\alpha}\widetilde{A}].$$

Skew information is defined by

$$I_{\alpha,\rho,\varepsilon}(A) = (1 + \varepsilon)\,\mathrm{Tr}[\rho\widetilde{A}^2] - \mathrm{Tr}[\rho^\alpha\widetilde{A}\rho^{1-\alpha}\widetilde{A}].$$

Theorem 4.2.8 assures the nonnegativity of the generalized skew information $I_{\alpha,\rho,\varepsilon}(A)$. Lemma 8.2.2 also assures the nonnegativity of the generalized skew correlation $\mathrm{Corr}_{\alpha,\rho,\varepsilon}(X, X)$ for any linear operator X. Then we have the following results. We omit their proofs. See [238] for the details.

Theorem 8.2.1 ([238]). *For any two self-adjoint operators A and B, any density operator ρ, any $0 < \alpha < 1$ and $\varepsilon \geq 0$, we have a generalized uncertainty relation:*

$$I_{\alpha,\rho,\varepsilon}(A)I_{\alpha,\rho,\varepsilon}(B) - \left|\mathrm{Re}\{\mathrm{Corr}_{\alpha,\rho,\varepsilon}(A, B)\}\right|^2 \geq \frac{\varepsilon^2}{4}\left|\mathrm{Tr}[\rho[A, B]]\right|^2.$$

We are interested in the relation of the left-hand side in Proposition 8.2.1 and Theorem 8.2.1. The following proposition gives the answer for this question.

Proposition 8.2.2 ([238]). *For any two self-adjoint operators A and B, any density operator ρ, any $0 < \alpha < 1$ and $\varepsilon \geq 0$, we have*

$$I_{\alpha,\rho,\varepsilon}(A)I_{\alpha,\rho,\varepsilon}(B) - \left|\mathrm{Re}\{\mathrm{Corr}_{\alpha,\rho,\varepsilon}(A, B)\}\right|^2 \geq \varepsilon^2 V_\rho(A)V_\rho(B) - \varepsilon^2\left|\mathrm{Re}\{\mathrm{Cov}_\rho(A, B)\}\right|^2.$$

Remark 8.2.1. Theorem 8.2.1 can be also proven by Proposition 8.2.1 and Proposition 8.2.2.

As an extension of [150, Theorem 2], we can prove the following inequality which give an affirmative answer to the conjecture in [149, p. 1572], as well. H. Kosaki also solved the same conjecture in [149, p. 1572] by another method [134].

Theorem 8.2.2 ([238]). *For any two self-adjoint operators A and B, any density operator ρ and any $0 < \alpha < 1$, we have*

$$V_\rho(A)V_\rho(B) - \left|\mathrm{Re}\{\mathrm{Cov}_\rho(A,B)\}\right|^2 \geq I_{\alpha,\rho,0}(A)I_{\alpha,\rho,0}(B) - \left|\mathrm{Re}\{\mathrm{Corr}_{\alpha,\rho,0}(A,B)\}\right|^2.$$

Remark 8.2.2. We point out that the inequality (8.2.3) given in [150, Theorem 1] is not true in general, because we have many counterexamples. For example, we set

$$\rho = \frac{1}{4}\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then we have $I_\rho(A)I_\rho(B) - |\mathrm{Re}\{\mathrm{Corr}_\rho(A,B)\}|^2 = \frac{7-4\sqrt{3}}{4}$ and $|\mathrm{Tr}[\rho[A,B]]|^2 = 1$. These imply

$$I_\rho(A)I_\rho(B) - \left|\mathrm{Re}\{\mathrm{Corr}_\rho(A,B)\}\right|^2 < \frac{1}{4}\left|\mathrm{Tr}[\rho[A,B]]\right|^2.$$

Thus the inequality (8.2.3) given in [150, Theorem 1] does not hold in general.

8.2.2 Schrödinger uncertainty relation with Wigner–Yanase skew information

We have the relation between $I_\rho(H)$ and $V_\rho(H)$ such that $0 \leq I_\rho(H) \leq V_\rho(H)$ so it is quite natural to consider that we have the further sharpened uncertainty relation for the Wigner–Yanase skew information:

$$I_\rho(A)I_\rho(B) \geq \frac{1}{4}\left|\mathrm{Tr}[\rho[A,B]]\right|^2.$$

However, the above relation failed (see [150, 134, 238].), as we have seen in Subsection 8.2.1. S. Luo then introduced the quantity $U_\rho(H)$ representing a quantum uncertainty excluding the classical mixture:

$$U_\rho(H) = \sqrt{V_\rho(H)^2 - \left(V_\rho(H) - I_\rho(H)\right)^2}, \tag{8.2.4}$$

and then he successfully showed a new Heisenberg-type uncertainty relation on $U_\rho(H)$ in [146]:

$$U_\rho(A)U_\rho(B) \geq \frac{1}{4}\left|\mathrm{Tr}[\rho[A,B]]\right|^2. \tag{8.2.5}$$

As stated in [146], the physical meaning of the quantity $U_\rho(H)$ can be interpreted as follows. For a mixed state ρ, the variance $V_\rho(H)$ has both classical mixture and quantum uncertainty. Also, the Wigner–Yanase skew information $I_\rho(H)$ represents a kind of quantum uncertainty [148, 147]. Thus, the difference $V_\rho(H) - I_\rho(H)$ has a classical mixture so we can regard that the quantity $U_\rho(H)$ has a quantum uncertainty excluding a classical mixture. Therefore, it is meaningful and suitable to study an uncertainty relation for a mixed state by use of the quantity $U_\rho(H)$.

K. Yanagi gave a one-parameter extension of the inequality (8.2.5) in [237], using the Wigner–Yanase–Dyson skew information $I_{\rho,\alpha}(H)$. Note that we have the following ordering among three quantities:

$$0 \le I_\rho(H) \le U_\rho(H) \le V_\rho(H). \tag{8.2.6}$$

The inequality (8.2.5) is a refinement of the original Heisenberg's uncertainty relation (8.2.1) in the sense of the above ordering (8.2.6).

In this subsection, we show the further strong inequality (Schrödinger-type uncertainty relation) for the quantity $U_\rho(H)$ representing a quantum uncertainty.

To show our main theorem, we prepare the definition for a few quantities and a lemma representing properties on their quantities.

Definition 8.2.2. For a quantum state ρ and an observable H, we define the following quantities:

(i) The Wigner–Yanase skew information:

$$I_\rho(H) = \frac{1}{2}\operatorname{Tr}\left[\left(i[\rho^{1/2}, H_0]\right)^2\right] = \operatorname{Tr}[\rho H_0^2] - \operatorname{Tr}[\rho^{1/2} H_0 \rho^{1/2} H_0],$$

where $H_0 = H - \operatorname{Tr}[\rho H]I$ and $[X, Y] = XY - YX$ is a commutator.

(ii) The quantity associated to the Wigner–Yanase skew information:

$$J_\rho(H) = \frac{1}{2}\operatorname{Tr}\left[\left(\{\rho^{1/2}, H_0\}\right)^2\right] = \operatorname{Tr}[\rho H_0^2] + \operatorname{Tr}[\rho^{1/2} H_0 \rho^{1/2} H_0],$$

where $\{X, Y\} = XY + YX$ is an **anticommutator**.

(iii) The quantity representing a quantum uncertainty:

$$U_\rho(H) = \sqrt{V_\rho(H)^2 - \left(V_\rho(H) - I_\rho(H)\right)^2}.$$

For two quantities $I_\rho(H)$ and $J_\rho(H)$, by simple calculations, we have

$$I_\rho(H) = \operatorname{Tr}[\rho H^2] - \operatorname{Tr}[\rho^{1/2} H \rho^{1/2} H]$$

and

$$\begin{aligned} J_\rho(H) &= \operatorname{Tr}[\rho H^2] + \operatorname{Tr}[\rho^{1/2} H \rho^{1/2} H] - 2\left(\operatorname{Tr}[\rho H]\right)^2 \\ &= 2V_\rho(H) - I_\rho(H), \end{aligned} \tag{8.2.7}$$

which implies $I_\rho(H) \le J_\rho(H)$. In addition, we have the following relations.

Lemma 8.2.3.

(i) *For a quantum state* ρ *and an observable* H, *we have the following relation among* $I_\rho(H), J_\rho(H)$ *and* $U_\rho(H)$:

$$U_\rho(H) = \sqrt{I_\rho(H)J_\rho(H)}.$$

(ii) *For a spectral decomposition of* $\rho = \sum_{j=1}^{\infty} \lambda_j |\phi_j\rangle\langle\phi_j|$, *putting* $h_{ij} = \langle\phi_i|H_0|\phi_j\rangle$, *we have*

$$I_\rho(H) = \sum_{i<j} (\sqrt{\lambda_i} - \sqrt{\lambda_j})^2 |h_{ij}|^2.$$

(iii) *For a spectral decomposition of* $\rho = \sum_{j=1}^{\infty} \lambda_j |\phi_j\rangle\langle\phi_j|$, *putting* $h_{ij} = \langle\phi_i|H_0|\phi_j\rangle$, *we have*

$$J_\rho(H) \geq \sum_{i<j} (\sqrt{\lambda_i} + \sqrt{\lambda_j})^2 |h_{ij}|^2.$$

The relation (i) immediately follows from (8.2.7). See [237] for the proofs of (ii) and (iii).

Theorem 8.2.3 ([62]). *For a quantum state (density operator)* ρ *and two observables (self-adjoint operators)* A *and* B, *we have*

$$U_\rho(A)U_\rho(B) - \left|\mathrm{Re}\{\mathrm{Corr}_\rho(A,B)\}\right|^2 \geq \frac{1}{4}\left|\mathrm{Tr}[\rho[A,B]]\right|^2, \tag{8.2.8}$$

where the correlation measure is defined by

$$\mathrm{Corr}_\rho(X,Y) = \mathrm{Tr}[\rho X^* Y] - \mathrm{Tr}[\rho^{1/2} X^* \rho^{1/2} Y]$$

for any operators X *and* Y.

Proof. We take a spectral decomposition $\rho = \sum_{j=1}^{\infty} \lambda_j |\phi_j\rangle\langle\phi_j|$. If we put $a_{ij} = \langle\phi_i|A_0|\phi_j\rangle$ and $b_{ji} = \langle\phi_j|B_0|\phi_i\rangle$, where $A_0 = A - \mathrm{Tr}[\rho A]I$ and $B_0 = B - \mathrm{Tr}[\rho B]I$, then we have

$$\begin{aligned}\mathrm{Corr}_\rho(A,B) &= \mathrm{Tr}[\rho AB] - \mathrm{Tr}[\rho^{1/2} A \rho^{1/2} B] = \mathrm{Tr}[\rho A_0 B_0] - \mathrm{Tr}[\rho^{1/2} A_0 \rho^{1/2} B_0] \\ &= \sum_{i,j=1}^{\infty} (\lambda_i - \lambda_i^{1/2}\lambda_j^{1/2}) a_{ij} b_{ji} = \sum_{i \neq j} (\lambda_i - \lambda_i^{1/2}\lambda_j^{1/2}) a_{ij} b_{ji} \\ &= \sum_{i<j} \{(\lambda_i - \lambda_i^{1/2}\lambda_j^{1/2}) a_{ij} b_{ji} + (\lambda_j - \lambda_j^{1/2}\lambda_i^{1/2}) a_{ji} b_{ij}\}.\end{aligned}$$

Thus we have

$$\left|\mathrm{Corr}_\rho(A,B)\right| \leq \sum_{i<j} \{\left|\lambda_i - \lambda_i^{1/2}\lambda_j^{1/2}\right| |a_{ij}||b_{ji}| + \left|\lambda_j - \lambda_j^{1/2}\lambda_i^{1/2}\right| |a_{ji}||b_{ij}|\}.$$

Since $|a_{ij}| = |a_{ji}|$ and $|b_{ij}| = |b_{ji}|$, taking a square of both sides and then using Schwarz inequality for a scalar and Lemma 8.2.3, we have

$$\begin{aligned}
\left|\mathrm{Corr}_\rho(A,B)\right|^2 &\le \left\{\sum_{i<j}\{|\lambda_i - \lambda_i^{1/2}\lambda_j^{1/2}| + |\lambda_j - \lambda_j^{1/2}\lambda_i^{1/2}|\}|a_{ij}||b_{ji}|\right\}^2 \\
&= \left\{\sum_{i<j}(\lambda_i^{1/2} + \lambda_j^{1/2})|\lambda_i^{1/2} - \lambda_j^{1/2}||a_{ij}||b_{ji}|\right\}^2 \\
&\le \left\{\sum_{i<j}(\sqrt{\lambda_i} - \sqrt{\lambda_j})^2|a_{ij}|^2\right\}\left\{\sum_{i<j}(\sqrt{\lambda_i} + \sqrt{\lambda_j})^2|b_{ij}|^2\right\} \\
&\le I_\rho(A)J_\rho(B).
\end{aligned}$$

By the similar way, we also have $|\mathrm{Corr}_\rho(A,B)|^2 \le I_\rho(B)J_\rho(A)$ Thus we have $|\mathrm{Corr}_\rho(A,B)|^2 \le U_\rho(A)U_\rho(B)$, which is equivalent to the inequality

$$U_\rho(A)U_\rho(B) - \left|\mathrm{Re}\{\mathrm{Corr}_\rho(A,B)\}\right|^2 \ge \frac{1}{4}\left|\mathrm{Tr}[\rho[A,B]]\right|^2,$$

since we have $|\mathrm{Im}\{\mathrm{Corr}_\rho(A,B)\}|^2 = \frac{1}{4}|\mathrm{Tr}[\rho[A,B]]|^2$. □

Theorem 8.2.3 improves the uncertainty relation (8.2.5) shown in [146], in the sense that the upper bound of the right-hand side of our inequality (8.2.8) is tighter than that of Luo's one (8.2.5).

Remark 8.2.3. For a pure state $\rho = |\varphi\rangle\langle\varphi|$, we have $I_\rho(H) = V_\rho(H)$ which implies $U_\rho(H) = V_\rho(H)$ for an observable H and $\mathrm{Corr}_\rho(A,B) = \mathrm{Cov}_\rho(A,B)$ for two observables A and B. Therefore, our Theorem 8.2.3 coincides with the Schrödinger uncertainty relation (8.2.2) for a particular case that a given quantum state is a pure state, $\rho = |\varphi\rangle\langle\varphi|$.

Remark 8.2.4. As a similar problem, we may consider the following uncertainty relation:

$$U_\rho(A)U_\rho(B) - \left|\mathrm{Re}\{\mathrm{Cov}_\rho(A,B)\}\right|^2 \ge \frac{1}{4}\left|\mathrm{Tr}[\rho[A,B]]\right|^2.$$

However, the above inequality does not hold in general, since we have a counterexample as follows. We take

$$\rho = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \quad A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and then we have

$$U_\rho(A)U_\rho(B) - \left|\mathrm{Re}\{\mathrm{Cov}_\rho(A,B)\}\right|^2 - \frac{1}{4}\left|\mathrm{Tr}[\rho[A,B]]\right|^2 = -\frac{3}{4}.$$

Remark 8.2.5. From Theorem 8.2.3 and Remark 8.2.4, we may expect that the following inequality holds:

$$\left|\mathrm{Re}\{\mathrm{Cov}_\rho(A,B)\}\right|^2 \geq \left|\mathrm{Re}\{\mathrm{Corr}_\rho(A,B)\}\right|^2. \tag{8.2.9}$$

However, the above inequality does not hold in general, since we have a counterexample as follows. We take

$$\rho = \frac{1}{10}\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}, \quad A = \begin{pmatrix} 4 & 4 \\ 4 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & -1 \\ -1 & 2 \end{pmatrix},$$

and then we have

$$\left|\mathrm{Re}\{\mathrm{Cov}_\rho(A,B)\}\right|^2 - \left|\mathrm{Re}\{\mathrm{Corr}_\rho(A,B)\}\right|^2 \approx -0.1539.$$

Actually, from Theorem 8.2.3, the example in Remark 8.2.4 and the above example, we find that there is no ordering between $|\mathrm{Re}\{\mathrm{Cov}_\rho(A,B)\}|^2$ and $|\mathrm{Re}\{\mathrm{Corr}_\rho(A,B)\}|^2$.

Remark 8.2.6. The example given in Remark 8.2.4 shows

$$V_\rho(A)V_\rho(B) - \left|\mathrm{Re}\{\mathrm{Cov}_\rho(A,B)\}\right|^2 - \left(U_\rho(A)U_\rho(B) - \left|\mathrm{Re}\{\mathrm{Corr}_\rho(A,B)\}\right|^2\right) \approx -0.232051.$$

The example given in Remark 8.2.5 also shows

$$V_\rho(A)V_\rho(B) - \left|\mathrm{Re}\{\mathrm{Cov}_\rho(A,B)\}\right|^2 - \left(U_\rho(A)U_\rho(B) - \left|\mathrm{Re}\{\mathrm{Corr}_\rho(A,B)\}\right|^2\right) \approx 13.7862.$$

Therefore, there is no ordering between $V_\rho(A)V_\rho(B) - |\mathrm{Re}\{\mathrm{Cov}_\rho(A,B)\}|^2$ and $U_\rho(A)U_\rho(B) - |\mathrm{Re}\{\mathrm{Corr}_\rho(A,B)\}|^2$ so we can conclude that neither the inequality (8.2.2) nor the inequality (8.2.8) is uniformly better than the other.

As we have seen, we proved a new Schrödinger-type uncertainty relation for a quantum state (generally a mixed state). Our result coincides with the original Schrödinger uncertainty relation for a particular case that a quantum state is a pure state. In addition, our result improves the uncertainty relation shown in [146] and as well as the original Heisenberg uncertainty relation. Later, this topic has been studied in more general setting with metric adjusted skew information. See [84, 239] and references therein.

Bibliography

[1] Aczél J. Some general methods in the theory of functional equations in one variable. New applications of functional equations. Usp Mat Nauk. 1956;11:3–68 (in Russian).

[2] Adamek M. On a problem connected with strongly convex functions. Math Inequal Appl. 2016;19(4):1287–93.

[3] Agarwal RP, Dragomir SS. A survey of Jensen type inequalities for functions of selfadjoint operators in Hilbert spaces. Comput Math Appl. 2010;59(12):3785–812.

[4] Alić M, Bullen PS, Pečarić JE, Volenec V. On the geometric–arithmetic mean inequality for matrices. Math Commun. 1997;2(2):125–8.

[5] Amari SI, Nagaoka H. Methods of information geometry. London: Oxford University Press; 2000.

[6] Ando T. Concavity of certain maps on positive definite matrices and applications to Hadamard products. Linear Algebra Appl. 1979;26:203–41.

[7] Ando T, Hiai F. Log-majorization and complementary Golden–Thompson type inequalities. Linear Algebra Appl. 1994;197/198:113–31.

[8] Ando T, Hiai F. Operator log-convex functions and operator means. Math Ann. 2011;350(3):611–30.

[9] Ando T, Hiai F, Okubo K. Trace inequalities for multiple products of two matrices. Math Inequal Appl. 2000;3:307–18.

[10] Ando T, Zhan X. Norm inequalities related to operator monotone functions. Math Ann. 1999;315:771–80.

[11] Araki H. On an inequality of Lieb and Thirring. Lett Math Phys. 1990;19:167–70.

[12] Audenaert KMR. A sharp continuity estimate for the von Neumann entropy. J Phys A. 2007;40:8127–36.

[13] Audenaert KMR et al.. Discriminating states: The quantum chernoff bound. Phys Rev Lett. 2007;98:160501.

[14] Audenaert KMR. An Araki-Lieb-Thirring inequality for geometrically concave and geometrically convex functions. Linear Algebra Appl. 2013;438:3454–62.

[15] Bebiano N, da Providência J Jr, Lemos R. Matrix inequalities in statistical mechanics. Linear Algebra Appl. 2004;376:265–73.

[16] Belavkin VP, Staszewski P. C^*-algebraic generalization of relative entropy and entropy. Ann Inst Henri Poincaré A. 1982;37:51–8.

[17] Bellman R. On an inequality concerning an indefinite form. Am Math Mon. 1956;63:108–9.

[18] Belmega E-V, Lasaulce S, Debbah M. A trace inequality for positive definite matrices. JIPAM J Inequal Pure Appl Math. 2009;10(1):5.

[19] Belmega EV, Jungers M, Lasaulce S. A generalization of a trace inequality for positive definite matrices. Aust J Math Anal Appl. 2011;7(2):26.

[20] Bhatia R. Matrix analysis. New York: Springer; 1997.

[21] Bhatia R. Interpolating the arithmetic–geometric mean inequality and its operator version. Linear Algebra Appl. 2006;413(2/3):355–63.

[22] Bhatia R. Positive definite matrices. Princeton: Princeton University Press; 2007.

[23] Bhatia R, Davis C. More matrix forms of the arithmetic–geometric mean inequality. SIAM J Matrix Anal Appl. 1993;14:132–6.

[24] Bhatia R, Davis C. More operator versions of the Schwarz inequality. Commun Math Phys. 2000;215:239–44.

[25] Bhatia R, Kittaneh F. Notes on matrix arithmetic–geometric mean inequalities. Linear Algebra Appl. 2000;308:203–11.

https://doi.org/10.1515/9783110643473-009

[26] Bhatia R, Kosaki H. Mean matrices and infinite divisibility. Linear Algebra Appl. 2007;424:36–54.
[27] Bobylev NA, Krasnoselsky MA. Extremum Analysis (degenerate cases). Moscow. 1981. Preprint, 52 pages (in Russian).
[28] Bourin JC, Hiai F. Jensen and Minkowski inequalities for operator means and anti-norms. Linear Algebra Appl. 2014;456:22–53.
[29] Bourin JC, Lee EY, Fujii M, Seo Y. A matrix reverse Hölder inequality. Linear Algebra Appl. 2009;431:2154–9.
[30] Bullen PS. A dictionary of inequalities. Reading: Addison Wesley Longman Limited; 1998.
[31] Bullen PS. Handbook of means and their inequalties. Norwell: Kluwer Academic Publishers; 2003.
[32] Carlen EA, Lieb EH. Remainder terms for some quantum entropy inequalities. J Math Phys. 2014;55:042201.
[33] Cheng KW. Majorization: Its extensions and preservation theorems. Tech Rep, No 121. Department of Statistics, Stanford University; 1977.
[34] Choi MD. A Schwarz inequality for positive linear maps on C^*-algebras. Ill J Math. 1974;18:565–74.
[35] Cloud MJ, Drachman BC, Lebedev LP. Inequalities with applications to engineering. Berlin: Springer; 1998.
[36] Cover TM, Thomas JA. Elements of information theory. 2nd ed. New York: Wiley and Sons; 2006.
[37] Cvetkovski Z. Inequalities, theorems, techniques and selected problems. Berlin: Springer; 2012.
[38] Davis C. A Schwarz inequality for convex operator functions. Proc Am Math Soc. 1957;8:42–4.
[39] Dragomir SS. A generalization of Aczél's inequality in inner product spaces. Acta Math Hung. 1994;65:141–8.
[40] Dragomir SS. Bounds for the normalized Jensen functional. Bull Aust Math Soc. 2006;74(3):471–6.
[41] Dragomir SS. Power inequalities for the numerical radius of a product of two operators in Hilbert spaces. Sarajevo J Math. 2009;5(18):269–78.
[42] Dragomir SS. On new refinements and reverse of Young's operator inequality. RGMIA Res Rep Collect. 2015;18:135.
[43] Dragomir SS. A note on Young's inequality. Rev R Acad Cienc Exactas Fís Nat, Ser A Mat. 2017;111:349–54.
[44] Dragomir SS, Buşe C. Refinements and reverses for the relative operator entropy $S(A|B)$ when $B \geq A$. RGMIA Res Rep Collect. 2016;19:168.
[45] Ebadian A, Nikoufar I, Gordji ME. Perspectives of matrix convex functions. Proc Natl Acad Sci USA. 2011;108:7313–4.
[46] Effros E, Hansen F. Non-commutative perspectives. Ann Funct Anal. 2014;5:74–9.
[47] Fu X, He C. Some operator inequalities for positive linear maps. Linear Multilinear Algebra. 2015;63:571–7.
[48] Fuchs L. A new proof of an inequality of Hardy–Littlewood–Pólya. Mat Tidssker. 1947;13:53–4.
[49] Fujii JI. A trace inequality arising from quantum information theory. Linear Algebra Appl. 2005;400:141–6.
[50] Fujii JI, Fujii M, Seo Y. An extension of the Kubo-Ando theory: Solidarities. Math Jpn. 1990;35(2):387–96.
[51] Fujii JI, Izumino S, Seo Y. Determinant for positive operators and Specht's theorem. Sci Math Jpn. 1998;1:307–10.

[52] Fujii JI, Kamei E. Relative operator entropy in non–commutative information theory. Math Jpn. 1989;34:341–8.

[53] Fujii JI, Nakamura M, Pečarić J, Seo Y. Bounds for the ratio and difference between parallel sum and series via Mond-Pečarić method. Math Inequal Appl. 2006;9(4):749–59.

[54] Fujii JI, Seo Y. Tsallis relative operator entropy with negative parameters. Adv Oper Theory. 2016;1(2):219–35.

[55] Fujii M, Furuichi S, Nakamoto R. Estimations of Heron means for positive operators. J Math Inequal. 2016;10(1):19–30.

[56] Fujii M, Furuta T, Kamei E. Furuta's inequality and its application to Ando's theorem. Linear Algebra Appl. 1993;179:161–9.

[57] Fujii M, Izumino S, Nakamoto R, Seo Y. Operator inequalities related to Cauchy–Schwarz and Hölder–McCarthy inequalities. Nihonkai Math J. 1997;8:117–22.

[58] Furuichi S. Trace inequalities in nonextensive statistical mechanics. Linear Algebra Appl. 2006;418:821–7.

[59] Furuichi S. A mathematical review of the generalized entropies and their matrix trace inequalities. In: Proceedings of WEC 2007. 2007. p. 840–5.

[60] Furuichi S. Matrix trace inequalities on the Tsallis entropies. J Inequal Pure Appl Math. 2008;9(1):1.

[61] Furuichi S. On the maximum entropy principle and the minimization of the Fisher information in Tsallis statistics. J Math Phys. 2009;50(1):013303.

[62] Furuichi S. Schrödinger uncertainty relation with Wigner–Yanase skew information. Phys Rev A. 2010;82:034101.

[63] Furuichi S. Inequalities for Tsallis relative entropy and generalized skew information. Linear Multilinear Algebra. 2011;59(10):1143–58.

[64] Furuichi S. On refined Young inequalities and reverse inequalities. J Math Inequal. 2011;5(1):21–31.

[65] Furuichi S. Refined Young inequalities with Specht's ratio. J Egypt Math Soc. 2012;20:46–9.

[66] Furuichi S. Operator inequalities among arithmetic mean, geometric mean and harmonic mean. J Math Inequal. 2014;8(3):669–72.

[67] Furuichi S. Unitarily invariant norm inequalities for some means. J Inequal Appl. 2014;2014:158.

[68] Furuichi S. Precise estimates of bounds on relative operator entropies. Math Inequal Appl. 2015;18(3):869–77.

[69] Furuichi S. Alternative proofs of the generalized reverse Young inequalities. Adv Inequal Appl. 2017;12:1–6.

[70] Furuichi S. A counter example for refined reverse Young inequality with Specht's ratio. Adv Inequal Appl. 2017;13:1–2.

[71] Furuichi S. Further improvements of Young inequality. Rev R Acad Cienc Exactas Fís Nat, Ser A Mat. 2019;113(1):255–66.

[72] Furuichi S. Note on constants appearing in refined Young inequalities. J Inequal Spec Funct. 2019;10(3):1–8.

[73] Furuichi S, Ghaemi MB, Gharakhanlu N. Generalized reverse Young and Heinz inequalities. Bull Malays Math Sci Soc. 2019;42(1):267–84.

[74] Furuichi S, Kuriyama K, Yanagi K. Trace inequalities for products of matrices. Linear Algebra Appl. 2009;430:2271–6.

[75] Furuichi S, Lin M. A matrix trace inequality and its application. Linear Algebra Appl. 2010;433:1324–8.

[76] Furuichi S, Lin M. Refinements of the trace inequality of Belmega, Lasaulce and Debbah. Aust J Math Anal Appl. 2011;7(2):23.

[77] Furuichi S, Minculute N. Alternative reverse inequalities for Young's inequality. J Math Inequal. 2011;5(4):595–600.

[78] Furuichi S, Minculete N. Inequalities for relative operator entropies and operator means. Acta Math Vietnam. 2018;43(4):607–18.

[79] Furuichi S, Moradi HR. Some refinements of classical inequalities. Rocky Mt J Math. 2018;48(7):2289–309.

[80] Furuichi S, Moradi HR. On further refinements for Young inequalities. Open Math. 2018;16:1478–782.

[81] Furuichi S, Moradi HR, Zardadi A. Some new Karamata type inequalities and their applications to some entropies. Rep Math Phys. 2019;84(2):201–14.

[82] Furuichi S, Moradi HR. New inequalities of the Kantorovich type with two negative parameters. arXiv:1710.02937.

[83] Furuichi S, Moradi HR, Sababheh M. New sharp inequalities for operator means. Linear Multilinear Algebra. 2019;67(8):1567–78.

[84] Furuichi S, Yanagi K. Schrödinger uncertainty relation, Wigner–Yanase–Dyson skew information and metric adjusted correlation measure. J Math Anal Appl. 2012;388:1147–56.

[85] Furuichi S, Yanagi K. Bounds of the logarithmic mean. J Inequal Appl. 2013;2013:535.

[86] Furuichi S, Yanagi K, Kuriyama K. Fundamental properties of Tsallis relative entropy. J Math Phys. 2004;45:4868–76.

[87] Furuichi S, Yanagi K, Kuriyama K. A note on operator inequalities of Tsallis relative operator entropy. Linear Algebra Appl. 2005;407:19–31.

[88] Furuichi S, Yanagi K, Kuriyama K. A generalized Fannes' inequality. JIPAM J Inequal Pure Appl Math. 2007;8(1):1–6.

[89] Furuichi S, Yanagi K, Kuriyama K. On bounds for symmetric divergence measures. AIP Conf Proc. 2017;1853:080002.

[90] Furuta T. Applications of order preserving operator inequalities. Oper Theory, Adv Appl. 1992;59:180–90.

[91] Furuta T. Operator inequalities associated with Hölder–McCarthy and Kantorovich inequalities. J Inequal Appl. 1998;2:137–48.

[92] Furuta T. Invitation to linear operators: From matrix to bounded linear operators on a Hilbert space. London: Taylor and Francis; 2002.

[93] Furuta T. Parametric extensions of Shannon inequality and its reverse one in Hilbert space operators. Linear Algebra Appl. 2004;381:219–35.

[94] Furuta T. Basic properties of the generalized Kantorovich constant $K(p) = \frac{h^p - h}{(p-1)(h-1)}\left(\frac{p-1}{p}\frac{h^p-1}{h^p-h}\right)^p$. Acta Sci Math. 2004;70:319–37.

[95] Furuta T. Reverse inequalities involving two relative operator entropies and two relative entropies. Linear Algebra Appl. 2005;403:24–30.

[96] Furuta T. Two reverse inequalities associated with Tsallis relative operator entropy via generalized Kantorovich constant and their applications. Linear Algebra Appl. 2006;412(2):526–37.

[97] Furuta T, Mićić J, Pečarić J, Seo Y. Mond–Pečarić method in operator inequalities. Series Monographs in Inequalities. vol. 1. Zagreb: Element; 2005.

[98] Garling DJH. Inequalities, A journey into linear analysis. Cambridge: Cambridge University Press; 2007.

[99] Ghaemi MB, Kaleibary V. Some inequalities involving operator monotone functions and operator means. Math Inequal Appl. 2016;19:757–64.

[100] Ghaemi MB, Kaleibary V, Furuichi S. On reverses of the Golden-Thompson type inequalities. J Math Inequal. 2018;12(2):315–23.

[101] Gibilisco P, Hansen F. An inequality for expectation of means of positive random variables. Ann Funct Anal. 2017;8(1):142–51.
[102] Golden S, Lower bounds for the Helmholtz function. Phys. Rev.. 1965;137:B1127–8.
[103] Gümüş IH. A note on a conjecture about Wielandt's inequality. Linear Multilinear Algebra. 2015;63(9):1909–13.
[104] Gümüş IH, Moradi HR, Sababheh M. More accurate operator means inequalities. J Math Anal Appl. 2018;465(1):267–80.
[105] Hansen F. Perspectives and completely positive maps. Ann Funct Anal. 2017;8(2):168–76.
[106] Hardy GH, Littlewood JE, Pólya G. Inequalities. Cambridge: Cambridge University Press; 1952.
[107] Heinz E. Beiträge zur Strüngstheorie der Spektrallegung. Math Ann. 1951;123:415–38.
[108] Heisenberg W. Über den anschaulichen Inhalt der quantummechanischen Kinematik und Mechanik. Z Phys. 1927;43:172–98.
[109] Hiai F. Matrix Analysis: Matrix Monotone Functions, Matrix Means, and Majorization. Interdiscip Inf Sci. 2010;16(2):139–248.
[110] Hiai F, Kosaki H. Comparison of various means for operators. J Funct Anal. 1999;163:300–23.
[111] Hiai F, Kosaki H. Means for matrices and comparison of their norms. Indiana Univ Math J. 1999;48:899–936.
[112] Hiai F, Kosaki H. Means of Hilbert Space Operators. Lecture Notes in Mathematics. vol. 1820. Berlin: Springer; 2003.
[113] Hiai F, Kosaki H, Petz D, Ruskai MB. Families of completely positive maps associated with monotone metrics. Linear Algebra Appl. 2013;439:1749–91.
[114] Hiai F, Petz D. The Golden-Thompson trace inequality is complemented. Linear Algebra Appl. 1993;181:153–85.
[115] Hiriart-Urruty JB, Lemaréchal C. Fundamentals of Convex Analysis. Berlin: Springer; 2001.
[116] Hirzallah O, Kittaneh F. Matrix Young inequalities for the Hilbert–Schmidt norm. Linear Algebra Appl. 2000;308:77–84.
[117] Hoa DT, Moslehian MS, Conde C, Zhang P. An extension of the Pólya-Szegö operator inequality. Expo Math. 2017;35(2):212–20.
[118] Horn RA, Johnson CR. Matrix analysis. London: Cambridge University Press; 1985.
[119] Huang K. Statistical Mechanics. New York: John Wiley and Sons; 1987.
[120] Isa H, Ito M, Kamei E, Toyama H, Watanabe M. Relative operator entropy, operator divergence and Shannon inequality. Sci Math Jpn. 2012;75:289–98.
[121] Isa H, Ito M, Kamei E, Toyama H, Watanabe M. Shannon type inequalities of a relative operator entropy including Tsallis and Rényi ones. Ann Funct Anal. 2015;6:289–300.
[122] Ivelić S, Matković A, Pečarić J. On a Jensen-Mercer operator inequality. Banach J Math Anal. 2011;5(1):19–28.
[123] Kaleibary V, Furuichi S. Reverses of operator Aczél inequality. C R Math Acad Sci Paris. 2018;356(5):475–81.
[124] Kantorovich LV. Functional analysis and applied mathematics. Usp Mat Nauk. 1948;3:89–185 (in Russian).
[125] Karamata J. Sur une inégalité relative aux fonctions convexes. Publ Mat Univ Belgrade. 1932;1(1):145–7.
[126] Kim YO, Fujii JI, Fujii M, Seo Y. Kantorovich type inequalities for the difference with two negative parameters. Sci Math Jpn. 2010;72(2):171–83.
[127] Kittaneh F. Norm inequalities for fractional powers of positive operators. Lett Math Phys. 1993;27:279–85.
[128] Kittaneh F. Notes on some inequalities for Hilbert space operators. Publ Res Inst Math Sci. 1988;24:283–93.

[129] Kittaneh F. A numerical radius inequality and an estimate for the numerical radius of the Frobenius companion matrix. Stud Math. 2003;158(1):11–7.
[130] Kittaneh F, Manasrah Y. Improved Young and Heinz inequalities for matrices. J Math Anal Appl. 2010;36:262–9.
[131] Kittaneh F, Manasrah Y. Reverse Young and Heinz inequalities for matrices. Linear Multilinear Algebra. 2011;59:1031–7.
[132] Klamkin MS. An inequality in powers. Math Mag. 1965;38:249–50.
[133] Kosaki H. Arithmetic–geometric mean and related inequalities for operators. J Funct Anal. 1998;156:429–51.
[134] Kosaki H. Matrix trace inequality related to uncertainty principle. Int J Math. 2005;16:629–46.
[135] Kosaki H. Positive definiteness of functions with applications to operator norm inequalities. Mem Am Math Soc. 2011;212:997.
[136] Krnić M, Lovričević N, Pečarić J. Multidimensional Jensen's operator on a Hilbert space and applications. Linear Algebra Appl. 2012;436(7):2583–96.
[137] Kubo F, Ando T. Means of positive operators. Math Ann. 1980;264:205–24.
[138] Lee EY. A matrix reverse Cauchy–Schwarz inequality. Linear Algebra Appl. 2009;430:805–10.
[139] Lenard A. Generalization of the Golden-Thompson inequality $\mathrm{Tr}(e^A e^B) \geq \mathrm{Tr}e^{!+B}$. Indiana Univ Math J. 1971;21:457–67.
[140] Liao W, Wu J, Zhao J. New versions of reverse Young and Heinz mean inequalities with the Kantorovich constant. Taiwan J Math. 2015;19(2):467–79.
[141] Lieb EH. Convex trace functions and the Wigner–Yanase-Dyson conjecture. Adv Math. 1973;11:267–88.
[142] Lin M. Squaring a reverse AM-GM inequality. Stud Math. 2013;215:187–94.
[143] Lin M. On an operator Kantorovich inequality for positive linear maps. J Math Anal Appl. 2013;402:127–32.
[144] Lin M, Wolkowicz H. An eigenvalue majorization inequality for positive semidefinite block matrices. Linear Multilinear Algebra. 2012;60:1365–8.
[145] Löwner C. Über monotone Matrixfunktionen. Math Z. 1934;38:177–216.
[146] Luo S. Heisenberg uncertainty relation for mixed states. Phys Rev A. 2005;72:042110.
[147] Luo S. Quantum versus classical uncertainty. Theor Math Phys. 2005;143:681–8.
[148] Luo S, Zhang Q. Informational distance on quantum-state space. Phys Rev A. 2004;69:032106.
[149] Luo S, Zhang Z. An information charactrization of Schrödinger's uncertainty principle. J Stat Phys. 2004;114:1557–76.
[150] Luo S, Zhang Q. On skew information. IEEE Trans Inf Theory. 2004;50(8):1778–82. and Correction to "On skew information". IEEE Trans Inf Theory. 2005;51:4432.
[151] Marshall AW, Olkin I. Inequalities: Theory of majorization and its applications. San Diego: Academic Press; 1979.
[152] Marshall AW, Olkin I. Inequalities for the trace function. Aequ Math. 1985;29:35–9.
[153] Marshall AW, Olkin I. Matrix versions of Cauchy and Kantorovich inequalities. Aequ Math. 1990;40:89–93.
[154] Matković A, Pečarić J, Perić I. A variant of Jensen's inequality of Mercer's type for operators with applications. Linear Algebra Appl. 2006;418:551–64.
[155] Matharu JS, Moslehian MS, Aujla JS. Eigenvalue extensions of Bohr's inequality. Linear Algebra Appl. 2011;435(2):270–6.
[156] McCarthy CA. c_p. Isr J Math. 1967;5:249–71.
[157] Mercer AM. A variant of Jensen's inequality. J Inequal Pure Appl Math. 2003;4(4):73.
[158] Merentes N, Nikodem K. Remarks on strongly convex functions. Aequ Math. 2010;80(1–2):193–9.

[159] Mićić J, Moradi HR, Furuichi S. Choi–Davis–Jensen's inequality without convexity. J Math Inequal. 2018;12(4):1075–85.

[160] Mićić J, Pavić Z, Pečarić J. Jensen's inequality for operators without operator convexity. Linear Algebra Appl. 2011;434(5):1228–37.

[161] Mićić J, Pečarić J. Order among power means of positive operators, II. Sci Math Jpn. 2010;71(1):93–109.

[162] Mićić J, Pečarić J, Seo Y. Complementary inequalities to inequalities of Jensen and Ando based on the Mond–Pečarić method. Linear Algebra Appl. 2000;318:87–108.

[163] Mićić J, Pečarić J, Seo Y. Function order of positive operators based on the Mond–Pečarić method. Linear Algebra Appl. 2003;360:15–34.

[164] Mićić J, Pečarić J, Seo Y, Tominaga M. Inequalities for positive linear maps on Hermitian matrices. Math Inequal Appl. 2000;3:559–91.

[165] Mitrinović DS, Pečarić J, Fink AM. Classical and new inequalities in analysis. Norwell: Kluwer Academic Publishers; 1993.

[166] Mitori FC. About the precision in Jensen–Steffensen inequality. An Univ Craiova, Ser Mat Inform. 2010;37:73–84.

[167] Mitori FC, Minculete N. On the Jensen functional and the strong convexity. Bull Malays Math Sci Soc. 2018;41(1):311–9.

[168] Mond B, Pečarić J. Convex inequalities in Hilbert space. Houst J Math. 1993;19(3):405–20.

[169] Mond B, Pečarić J. Remarks on Jensen's inequality for operator convex functions. Ann Univ Mariae Curie-Skłodowska, Sect A. 1993;47(10):96–103.

[170] Mond B, Pečarič J. A matrix version of the Ky Fan generalization of the Kantorovich inequality. Linear Multilinear Algebra. 1994;36:217–21.

[171] Mond B, Pečarić J. Converses of Jensen's inequality for several operators. Rev Anal Numér Théor Approx. 1994;23:179–83.

[172] Mond B, Pečarić J. Operator convex functions of several variables. Soochow J Math. 1998;24:239–54.

[173] Moradi HR, Furuichi S. Improvement and generalization of some Jensen–Mercer-type inequalities. To appear in J Math Inequal.

[174] Moradi HR, Furuichi S. Two refinements for the operator Jensen–Mercer type inequality. arXiv:1709.01808.

[175] Moradi HR, Furuichi S. A complementary inequality to the information monotonicity for Tsallis relative operator entropy. Linear Multilinear Algebra. 2018. https://doi.org/10.1080/03081087.2018.1549532.

[176] Moradi HR, Furuichi S, Minculete N. Estimates for Tsallis relative operator entropy. Math Inequal Appl. 2017;20(4):1079–88.

[177] Moradi HR, Furuichi S, Heydarbeygi Z. New refinement of the operator Kantorovich inequality. Mathematics. 2019;7(2):139.

[178] Moradi HR, Furuichi S, Mitroi FC, Naseri R. An extension of Jensen's operator inequality and its application to Young inequality. Rev R Acad Cienc Exactas Fís Nat, Ser A Mat. 2019;113(2):605–14.

[179] Moradi HR, Furuichi S, Sababheh M. A better bound on the Jensen's operator inequality. arXiv:1806.09831.

[180] Moradi HR, Gümüş IH, Heydarbeygi Z. A glimpse at the operator Kantorovich inequality. Linear Multilinear Algebra. 2019;67(5):1031–6.

[181] Moradi HR, Omidvar ME. Complementary inequalities to improved AM-GM inequality. Acta Math Sin Engl Ser. 2017;33(12):1609–16.

[182] Moradi HR, Omidvar ME, Gümüş IH, Naseri R. A note on some inequalities for positive linear maps. Linear Multilinear Algebra. 2018;66(7):1449–60.

[183] Moradi HR, Omidvar ME, Adil Khan M, Nikodem K. Around Jensen's inequality for strongly convex functions. Aequ Math. 2018;92(1):25–37.
[184] Moslehian MS. Operator Aczél inequality. Linear Algebra Appl. 2011;434:1981–7.
[185] Moslehian MS, Mirzapour F, Morassaei A. Operator entropy inequalities. Colloq Math. 2013;2(130):159–68.
[186] Moslehian MS, Najafi H. An extension of Löwner-Heinz inequality. Linear Algebra Appl. 2012;437:2359–65.
[187] Moslehian MS, Nakamoto R, Seo Y. A Diaz-Metcalf type inequality for positive linear maps and its applications. Electron J Linear Algebra. 2011;22:179–90.
[188] Mosonyi M, Ogawa T. Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies. Commun Math Phys. 2015;334:1617–48.
[189] Nakamoto R, Nakamura M. Operator mean and Kantorovich inequality. Math Jpn. 1996;44(3):495–8.
[190] Nelson M, Nikodem K, Rivas S. Remarks on strongly Wright–convex functions. Ann Pol Math. 2011;102(3):271–8.
[191] von Neumann J. Thermodynamik quantenmechanischer Gesamtheiten. Gött Nachr. 1927;1927:273–91.
[192] Niculescu CP. Convexity according to the geometric mean. Math Inequal Appl. 2000;3(2):155–67.
[193] Niculescu CP, Persson LE. Old and new on the Hermite–Hadamard inequality. Real Anal Exch. 2004;29(2):663–86.
[194] Niculescu CP, Persson LE. Convex functions and their applications. 2nd ed. Berlin: Springer; 2018.
[195] Nielsen MA, Chuang IL. Quantum computation and quantum information. Cambridge: Cambridge University Press; 2000.
[196] Niezgoda M. A generalization of Mercer's result on convex functions. Nonlinear Anal. 2009;71(7):2771–9.
[197] Nikoufar I. On operator inequalities of some relative operator entropies. Adv Math. 2014;259:376–83.
[198] Olson MP. The selfadjoint operators of a von Neumann algebra from a conditionally complete lattice. Proc Am Math Soc. 1971;28:537–44.
[199] Ohya M, Petz D. Quantum entropy and its use. 2nd ed. Berlin: Springer; 2004.
[200] Pečarić J, Mićić J. Some functions reversing the order of positive operators. Linear Algebra Appl. 2005;396:175–87.
[201] Petz D. A variational expression for the relative entropy of states of a von Neumann algebra. Commun Math Phys. 1988;114:345–9.
[202] Petz D. Quantum information theory and quantum statistics. Berlin: Springer; 2004.
[203] Polyak BT. Existence theorems and convergence of minimizing sequences in extremum problems with restrictions. Sov Math Dokl. 1966;7:72–5.
[204] Popoviciu T. On an inequality. Gaz Mat Fiz, Ser A. 1959;11:451–61 (in Romanian).
[205] Rényi A. On measures of entropy and information. In: Proc. Fourth Berkeley, 1960. Symp Math Stat Prob. vol. 1. Berkeley: University of California Press; 1961. p. 547.
[206] Robertson HP. The uncertainty principle. Phys Rev. 1929;34:163–4.
[207] Ruskai MB, Stillinger FM. Convexity inequalities for estimating free energy and relative entropy. J Phys A. 1990;23:2421–37.
[208] Sababheh M. Interpolated inequalities for unitarily invariant norms. Linear Algebra Appl. 2015;475:240–50.
[209] Sababheh M. Convexity and matrix means. Linear Algebra Appl. 2016;506:588–602.

[210] Sababheh M. Log and harmonically log-convex functions related to matrix norms. Oper Matrices. 2016;10(2):453–65.
[211] Sababheh M. Improved Jensen's inequality inequality. Math Inequal Appl. 2017;20(2):389–403.
[212] Sababheh M. Means refinements via convexity. Mediterr J Math. 2017;14(3):125.
[213] Sababheh M, Choi D. A complete refinement of Young's inequality. J Math Anal Appl. 2016;440:379–93.
[214] Sababheh M, Moradi HR, Furuichi S. Exponential inequalities for positive linear mappings. J Funct Spaces. 2018;2018:54674131.
[215] Sababheh M, Moradi HR, Furuichi S. Operator inequalities via geometric convexity. Math Inequal Appl. 2019;22(4):1215–31.
[216] Sababheh M, Moradi HR, Furuichi S. Reversing Bellman operator inequality. To appear in J Math Inequal.
[217] Sason I. Tight Bounds for Symmetric Divergence Measures and a Refined Bound for Lossless Source Coding. IEEE Trans Inf Theory. 2015;61:701–7.
[218] Schrödinger E. About Heisenberg uncertainty relation. In: Proc Prussian Acad Sci Phys Math Sect. vol. XIX. 1930. p. 293.
[219] Seo Y. Reverses of the Golden–Thompson type inequalities due to Ando–Hiai–Petz. Banach J Math Anal. 2008;2:140–9.
[220] Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948;27:379–423 and 623–656.
[221] Shebrawi K, Albadawi H. Numerical radius and operator norm inequalities. J Inequal Appl. 2009;2009:1.
[222] Sheybani S, Omidvar ME, Moradi HR. New inequalities for operator concave functions involving positive linear maps. Math Inequal Appl. 2018;21(4):1167–74.
[223] Specht W. Zur Theorie der elementaren Mittel. Math Z. 1960;74:91–8.
[224] Thompson CJ. Inequality with applications in statistical mechanics. J Math Phys. 1965;6:1812–3.
[225] Thompson CJ. Inequalities and partial orders on matrix spaces. Indiana Univ Math J. 1971;21:469–80.
[226] Tominaga M. Specht's ratio in the Young inequality. Sci Math Jpn. 2001;5:525–30.
[227] Turkmen R, Paksoy VE, Zhang F. Some inequalities of majorization type. Linear Algebra Appl. 2012;437:1305–16.
[228] Tsallis C. Possible generalization of Bolzmann-Gibbs statistics. J Stat Phys. 1988;52:479–87.
[229] Umegaki H. Conditional expectation in an operator algebra, IV (entropy and information). Kodai Math Semin Rep. 1962;14:59–85.
[230] Wada S. Some ways of constructing Furuta–type inequalities. Linear Algebra Appl. 2014;457:276–86.
[231] Wada S. When does Ando-Hiai inequality hold? Linear Algebra Appl. 2018;540:234–43.
[232] Wang BY, Zhang F. Trace and eigenvalue inequalities for ordinary and Hadamard products of positive semidefinite Hermitian matrices. SIAM J Matrix Anal Appl. 1995;16(4):1173–83.
[233] Wigner EP, Yanase MM. Information content of distribution. Proc Natl Acad Sci USA. 1963;49:910–8.
[234] Wigner EP, Yanase MM. On the positive semidefinite nature of certain matrix expression. Can J Math. 1964;16:397–406.
[235] Wu J, Zhao J. Operator inequalities and reverse inequalities related to the Kittaneh-Manasrah inequalities. Linear Multilinear Algebra. 2014;62:884–94.
[236] Xue J, Hu X. Some generalizations of operator inequalities for positive linear maps. J Inequal Appl. 2016;2016:27.

[237] Yanagi K. Uncertainty relation on Wigner–Yanase–Dyson skew information. J Math Anal Appl. 2010;365:12–8.

[238] Yanagi K, Furuichi S, Kuriyama K. A generalized skew information and uncertainty relation. IEEE Trans Inf Theory. 2005;51:4401–4.

[239] Yanagi K, Furuichi S, Kuriyama K. Uncertainty relations for generalized metric adjusted skew information and generalized metric adjusted correlation measure. J Uncertain Anal Appl. 2013;1(12):1–14.

[240] Yanagi K, Kuriyama K, Furuichi S. Generalized Shannon inequalities based on Tsallis relative operator entropy. Linear Algebra Appl. 2005;394:109–18.

[241] Zhan X. Matrix inequalities. Berlin: Springer; 2002.

[242] Zhao X, Li L, Zuo H. Operator iteration on the Young inequality. J Inequal Appl. 2016;2016:302.

[243] Zhu L. New refinements of Young's inequality. Rev R Acad Cienc Exactas Fís Nat, Ser A, Math. 2019;113(2):909–15.

[244] Zlobec S. Characterization of convexifiable functions. Optimization. 2006;55(3):251–61.

[245] Zou L. Matrix versions of the classical Pólya inequality. ScienceAsia. 2013;39:204–7.

[246] Zou L. Operator inequalities associated with Tsallis relative operator entropy. Math Inequal Appl. 2015;18(2):401–6.

[247] Zou L, Jiang Y. Improved arithmetic–geometric mean inequality and its application. J Math Inequal. 2015;9:107–11.

[248] Zuo H, Shi G, Fujii M. Refined Young inequality with Kantorovich constant. J Math Inequal. 2011;5:551–6.

Index